Study of Protein structure by Electron Microscopy and Single Particle Image Processing

Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

Der Universität Basel

von

Venkata Prasad Dandey

von Indien

Basel, Schweiz, 2015

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät
Der Universität Basel
auf Antrag von

Prof. Dr.Henning Stahlberg(Biozentrum, Universität Basel), Fakultätsverantwortlicher
Prof. Dr.Volker Roth (Departement Informatik, Universität Basel), Koreferent

Basel,den 09.12.2014

Dr.Jörg Schibler (Dekan)

Abstract

The molecular assemblies form an important role in the understanding the machinery of living cells. These assemblies with thickness less than 1000nm and exhibiting different shapes, size and biochemical states are studied using Transmission Electron Microscopy (TEM). The ordered arrays of assemblies such as 2D crystals of membrane proteins and Helices are first reconstructed to near atomic resolution by image processing of its EM images. High-resolution structures of 2D crystals and helices using EM are very rare, due to inability of many complexes to form highly ordered arrays. If the proteins are large enough, it can be prepared as single particles and can be studied using EM. Single particle based approach is widely applicable approach for any symmetry and in the range of 0.1 - 100 MDa molecular weight complexes. This approach has caught up crystallography to reach near atomic resolution after the advancements in electron detector. In single particle approach many identical particles are expected to present on an EM grid and then 3D structure is calculated from the many views of the same molecule. The quality of the 3D structure is highly dependent on homogenous preparations of single particles.

The aim of this thesis is to determine the structure of molecular assemblies through single particle approach, focusing on their image processing. Also, the single particle based helical image processing techniques, applied to very thin and weakly ordered helical filaments, are also studied. In the methods section (Chapter 2-3), both methods related to conventional single particle reconstruction (SPR) for non-homogenous single particles and single particle based Iterative helical real space reconstruction (IHRSR) approaches are presented. The workflow of semi automated tool chain for SPR applied to heterogeneous sample, from automated particle picking to the final 3D density maps, is explained. The limitation of IHRSR technique is its dependence on initial guess helical symmetry parameters for the helical filaments. So, determining the helical symmetry by analysis of diffraction the pattern of the helical objects is explained in detail.
The second part of the thesis is the application of these methods in structural studies of unknown molecular assemblies. The SPR techniques are applied to study the heterogeneity of the Silent Information Regulatory (SIR) protein complexes. The flexibility among the different subunits of the SIR complex and existence of mixed structures are revealed using single particle based random conical tilt reconstruction (RCT) approach, which supports the results of MALDI-TOF mass spectrometer results. To analyze this sample, a semi-automated EM tool chain has been developed in house to automate Tecnai T12 electron microscope for collecting thousands of images per day and followed by semi automated SPR techniques.The structural studies of Pyrin domain of Apoptisis Speck like CARD domain (ASC-PYD) by Cryo-EM using IHRSR technique revealed the structure of its filaments to 3.7Å. A novel hybrid approach combining the solution NMR, solid-state NMR and Cryo-EM techniques are used to obtain precise structural

details of PYD filaments to atomic resolution. This novel method is expected to become a routine to complement both Cryo-EM and NMR results. In Chapter 6-7, the semi-automatic image acquisition and analysis methods used for TEM investigations in combination with isolation methods are used to extract the quantitative information on protein abundances.

Acknowledgements

Firstly, I would like to start by thanking my supervisor **Henning Stahlberg** for giving me the opportunity to work in his world's best equipped lab and advising me through out my Ph.D. He gave me the freedom to explore in the field of electron microscopy and automation, which helped me to move ahead in my career.

I would like to convey special gratitude to my co-supervisor **Volker Roth** and the chair **Thomas Vetter** for supporting me with scientific advice and feedback at different levels of my research.

Also, I am thankful to **Sebastian Hiller** and **Sussan Gasser** for collaborating with our lab and giving me the oppurtunity to work in their projects. It was nice working with **Lorenzo Sborgi** and **Stephanie Kueng** during this collaborations.

My fruitful results wouldn't have been possible without **Mohamed Chami** for his unconditional support and encouragement.

The contribution of **Dominic Giss** and his supervisor **Thomas Braun** for applying my work of automating T12 Electron Microscope for their quantitative TEM project ended up giving fruitful results.

Bill Anderson, you are the best tinker of Electron Microscope who saved me from all troubles. You showered me with your magic dust of knowledge, which helped me to successfully finish my automation projects. I should thank **Alexandra Graff Meyer** for introducing me to the world of electron microscopy at the early stage of my Ph.D. Kenny Goldie, with out you managing the lab equipped with a lot of sophisticated machines, would have been a nightmare to operate.

I also enjoyed the continuous scientific discussions and support from **Sebastian Scherer, Phillip Ringler** and especially with **Shirley Mueller**, who is a great asset to this lab.

Contents

Acknowledgements iv

Contents v

List of Figures viii

List of Tables x

1 Introduction 1
- 1.1 Macromolecules and their structure . . . 1
- 1.2 Heterogeneity in 2D and 3D . . . 2
  - 1.2.1 Sources of Heterogeneity . . . 2
  - 1.2.2 Methods for Computational Classification of Mixed Structures . . . 2
  - 1.2.3 Random Conical Tilt Reconstruction . . . 3
- 1.3 Algorithmic procedure for single particle reconstruction . . . 4
- 1.4 High resolution structure of helical polymers . . . 4
- 1.5 Structure and Aim of Dissertation . . . 5
  - 1.5.1 High-resolution structure of PYD filaments . . . 5
  - 1.5.2 EM Analysis of the basic units of the SIR complex . . . 6

2 Method I - Visualize the Heterogeneity by 3D Electron Microscopy 7
- 2.1 Abstract . . . 7
- 2.2 Introduction . . . 7
- 2.3 Class Averaging Techniques . . . 9
  - 2.3.1 Iterative reference alignment . . . 10
  - 2.3.2 Contrast transfer function estimation and correction . . . 12
- 2.4 Random Conical Reconstruction . . . 14
- 2.5 Tilt pair particle picking . . . 15
- 2.6 Processing of untitled images . . . 16
- 2.7 3-D reconstruction by Radon inversion . . . 16

3 Method II - Processing Images of Helical Structures 18
- 3.1 Introduction . . . 18
- 3.2 Real Space reconstruction of helical filaments . . . 19
  - 3.2.1 Intrinsic Ambiguities in indexing diffraction of helical filaments . . . 22

4 Sir4 dimerization promotes oligomerization of the SIR complex 24
4.1 Introduction to SIR complex . . . 25
4.2 Sample preparation . . . 26
4.3 Grid preparation . . . 26
4.4 MALDI-TOF analysis . . . 28
4.5 Data Collection . . . 28
4.5.1 Imaging and checking image quality with CM1 . . . 28
4.5.2 Automated Image acquisition using T12 microscope . . . 28
4.6 EM analysis of the basic units of the SIR complex . . . 30
4.7 Random Conical Tilt Reconstruction with Tilt pairs Sir 24 mutant complex (For heterogeneity analysis) . . . 30
4.7.1 Data collection of Tilt pairs . . . 30
4.7.2 2D and 3D Analysis . . . 33
4.7.3 Computing RCT reconstruction form each class average . . . 34
4.8 Conclusion from EM Analysis of the SIR complex basic units . . . 36

5 Structure and assembly of the mouse ASC filament by combined NMR spectroscopy and cryo-electron microscopy 38
5.1 Introduction . . . 39
5.2 Results . . . 42
5.2.1 ASC filament reconstitution *in vitro* . . . 42
5.2.2 Solid-state NMR spectroscopy of the mouse ASC-PYD filament . 43
5.2.3 Cryo-electron microscopy of the ASC-PYD filament . . . 43
5.2.4 Combined structure calculation of the mouse ASC-PYD filament 44
5.2.5 Structure of the ASC filament . . . 44
5.2.6 Assembly of ASC-PYD into the filament form . . . 45
5.2.7 Structural conservation of the ASC-PYD filament . . . 46
5.2.8 Dynamic and flexibility of the CARD domain . . . 47
5.3 Discussion . . . 49
5.4 Experimental Procedures . . . 51
5.4.1 Cloning, expression and purification of ASC-FL and ASC-PYD . 51
5.4.2 ASC-FL and ASC-PYD filament formation in vitro . . . 52
5.4.3 Solution NMR spectroscopy . . . 52
5.4.4 Solid-state NMR spectroscopy . . . 52
5.4.5 Cryo-EM microscopy and image reconstruction . . . 52
5.4.6 Structure calculation procedure . . . 53
5.4.7 Cell culture . . . 54
5.4.8 Retroviral Constructs . . . 54
5.4.9 Cell lines . . . 55
5.4.10 Inflammasome activation . . . 55
5.4.11 Confocal microscopy . . . 55

6 6. Exploring the interactome: Microfluidic isolation of proteins and interacting partners for visual analysis by quantitative electron microscopy 66
6.1 Introduction . . . 67
6.2 Experimental section . . . 68

6.2.1 Working principle . . . 68
6.2.2 Separation setup . . . 69
6.2.3 Loading of magnetic beads . . . 69
6.2.4 Antibody biotinylation . . . 70
6.2.5 Extraction and purification of target structures . . . 70
6.2.6 Recovery of target structures . . . 71
6.2.7 Washing procedure . . . 71
6.2.8 Cell preparation . . . 71
6.2.9 TEM grid preparation . . . 72
6.2.10 Image acquisition and processing . . . 72
6.2.11 Results and discussion . . . 73
6.3 Conclusions . . . 81
6.4 Supporting Information . . . 82
6.4.1 Photograph of the experimental setup . . . 82
6.4.2 Electron micrograph and 2D class averages of apoferritin . . . 83
6.4.3 Electron micrograph and 2D class averages of apoferritin . . . 84
6.4.4 Factors determining the signal transfer function of the method presented . . . 85
6.4.4.1 Signal transfer function of apoferritin to TEM grids . . . 86
6.4.5 Isolation of endogenous 20S proteasome in PBS buffer . . . 87
6.4.6 Electron micrograph and 2D class averages of 20S proteasomes . . . 88
6.4.7 COMSOL simulation of temperature rise in capillary and buffer . . . 88

7 Microfluidics to isolate untagged proteins from cell extracts for visual analysis by electron microscopy 90
7.1 Introduction . . . 91
7.2 Functional principle . . . 91
7.3 Experimental . . . 92
7.4 Results and discussion . . . 93
7.5 Conclusion . . . 95

8 Conclusion 96

List of Figures

2.1 The constrast transfer function . . . 13
2.2 Illustration of RCT technique . . . 15

3.1 Basic types of regular structures . . . 18
3.2 Iterative Helical Real Space Reconstruction . . . 19
3.3 Convergence plot of IHRSR algorithm for helical rise . . . 21
3.4 Convergence plot of IHRSR algorithm for angle between sub units $\Delta\phi$ . . . 21
3.5 Intrinsic Ambiguities in indexing diffraction of helical filaments . . . 22

4.1 Schematic of Sir3 protein . . . 27
4.2 Schematic diagram explaining gene silencing . . . 27
4.3 Co-expression of Sir4-Cm and Sir4-SII . . . 29
4.4 MALDI-TOF results for Sir2-4 mutant complex . . . 29
4.5 Semi Automated pipeline . . . 29
4.6 TEM analysis of negatively stained Sir24WT sample . . . 31
4.7 TEM analysis of negatively stained ir24cc sample. . . . 31
4.8 TEM analysis of negatively stained Sir3WT . . . 31
4.9 Full averages of EM analysis of Sir2-Sir4 hetero dimer, hetero tetramers and of Sir3 dimers. . . . 32
4.10 Gui for Tilt pair picking tool in EMAN2 package . . . 33
4.11 RCT analysis of SIR24 mutant complex . . . 35
4.12 Final RCT 3-D reconstruction after 3-D alignment . . . 36
4.13 EM analysis of Sir24 complexes and Sir3 homo dimer . . . 37

5.1 Structural determinants of the ASC filament formation. . . . 56
5.2 Structure determination of the ASC-PYD filament by a combination of ssNMR and cryo-EM. . . . 57
5.3 Three-dimensional assembly of the ASC-PYD filament. (A–B) . . . 58
5.4 Characterization of the ASC-PYD subunit interaction . . . 59
5.5 Structural comparison of mouse and human ASC filaments and evolutionary conservation. . . . 59
5.6 The CARD domain in the mouse ASC-FL filament is flexibly unfolded. 60
5.7 Effect of single point mutations on ASC-dependent signalling . . . 61
5.8 Sequence-specific NMR resonance assignments of ASC-PYD and ASC-FL filaments . . . 62
5.9 Contribution of cryo-Em restraints to the ASC-PYD structure calculation 63
5.10 Sequence-specific resonance assignments of monomeric ASC-PYD and comparison with the ASC-PYD fibril form. . . . 64

5.11 Comparison of secondary structure of the mouse and human monomeric ASC-PYD. 64

6.1 Principle of the microfluidic affinity isolation method. 74
6.2 Semi-automatic TEM procedure for image acquisition and data-analysis. 75
6.3 Affinity extraction and recovery of apoferritin 76
6.4 Signal transfer function of the microfluidic affinity isolation method for AF determined by qTEM. 77
6.5 Isolation of endogenous protein complexes 79
6.6 The detection of extracted protein binding partners by interaction-labeling. 80
6.7 Photograph of the experimental setup and higher magnification photograph of a magnetic bead plug (inset). 82
6.8 Negative stain transmission electron microscopy (TEM) of apoferritin (AF). 83
6.9 Negative stain electron micrograph of isolated 20S proteasomes 84
6.10 Factors determining the signal transfer function (SiTF) of the presented isolation method. 85
6.11 Signal transfer function (SiTF) of apoferritin (AF) particles to transmission electron microscopy (TEM) grids. 86
6.12 Isolation of endogenous 20S proteasome in PBS buffer. 87
6.13 Negative stain transmission electron microscopy (TEM) of purified 20S proteasomes. 88
6.14 Finite element simulations using COMSOL 89

7.1 Working principle of the method. A: Scheme of the composite material used for protein extraction. 92
7.2 Experimental setup used for the experiments. Inset: higher magnification showing a magnetic bead plug. 93
7.3 Comparison of contaminant background of photocleavage and competitive elution by biotin and imidazole. 93
7.4 Effect of illumination time on the recovery of protein. 94

List of Tables

5.1 Supplementary Table 1. Structural statistics for the mouse ASC PYD filament 65

Dedicated to my loving wife Silpa, my parents, In-laws and my sister Mallika.. . .

Chapter 1

Introduction

1.1 Macromolecules and their structure

A central fact of modern biology is that large polymeric molecules and their assemblies (proteins, DNA etc.) are vital to cell function. These macromolecules not only make biochemical cell reactions possible, but they also maintain cell structure, cause cell motion, sense and response to signals in the environment, etc. Macromolecules are able to do all this because of their three-dimensional structure. That allows then to dock with other molecules in order to facilitate reactions between them. In addition, structural flexibility allows motion, signal transduction, etc. To reconstruct the 3D structure of the macromolecules to atomic resolution, X-ray crystallography and Nuclear Magnetic Resonance were the only two prominent methods until the invention of Direct Electron Detectors (DDD) in cryogenic Electron Microscopy (cryo-EM) that made is possible for most of the macromolecular structures to be reconstructed to near atomic resolution ($\sim$ 3Å). In cryo-EM, random projections of the macromolecule are obtained and are used to reconstruct three dimensional structure of the macromolecule.
In Electron Microscopy the major applications concerning to biology are the interpretation of the images on a qualitative and descriptive level. For a descriptive level there is a need for an accurate three-dimensional (3D) representation, which reveals its interior density variations of the structures. So, the role of 3D electron microscopy is to visualize complex biological structures. In order to bridge the gap of several orders of magnitude between X-ray crystallography and light microscopy.
In image processing of electron microscopy, the particles are commonly referred to biological macromolecules and their assemblies. One among the three techniques to reconstruct a 3D structure from EM images is Single-Particle Reconstruction (SPR). The term cryo-EM will refer to the experimental process of obtaining images, while SPR will

refer to the algorithmic problem of reconstructing the three-dimensional structure from the projections. Helical filaments were the first structures to be reconstructed in three dimensions from electron microscopy images. Biological macromolecules which form helical structures are generally discrete and the discrete points at regular intervals where the subunits are located and it is referred as particles and so, modified SPR technique is used as a powerful method to reconstruct helical structures.

1.2 Heterogeneity in 2D and 3D

1.2.1 Sources of Heterogeneity

Heterogeneity in the macromolecular structure limits the resolution of the reconstructed structure by EM due to conformational changes, which complicates the image processing. During the sample preparation by vitrification can freeze the different confirmations of the biological complex and these will reflect in the images [1]. There are several sources that can result in sample heterogeneity, few of them are: (i) incomplete coverage of ligand in molecular complex, [2] (ii) structural transformation along the intermediate states [3–6] and (iii) variation in symmetry and size within the multiple oligomeric states. Biochemical Stabilization or trapping distinct conformation and/or Gradient fixation can partially or fully avoid heterogeneity. But cannot cannot be eliminated in all the cases [7].

1.2.2 Methods for Computational Classification of Mixed Structures

To sort mixed structures using computational sorting methods, the first approach is to recognize the heterogeneity of the sample in 2D before obtaining the 3D structure [7]. In this, "apriori" methods are mainly based on Multivariate Statistical Analysis(MSA) applied to the features of the 2D images to find structural variations. The 2D analysis will help in differentiating the structural variations and orientation differences. Multi stage MSA and classification can be applied to sort major variations, which is replicated in lower order Eigen images. For example, a technique to sort variable occupancy of a substrate from a data set is to apply two-stage approach, first to classify global variance due to orientation differences and then second classification of localized variations produced due to substrate binding. This process does not require calculating the angular orientation of the particles using 3D model and it has been successful to discriminate even the 5% variations in total size of particle like ribosome_EF-G complexes [5, 8–11].

The third approach is heterogeneity analysis by 3D reconstructions. Classification of the dataset into a group of 3D reconstructions reveals the variations in 3D that is more accurate than 2D analysis. Many 3D models are calculated from 2D classification and averages for variance analysis in 3D. In [12], a fast and efficient approach to calculate the 3D variance by the technique called "bootstrap" is introduced. In this method, heterogeneous data sets are resampled into subsets of randomly selected images. These images are randomly sampled and may be used more than once in any of the subset groups. Spatial orientations of each image of the resampled sets are already assigned with the initial 3D map (consensus model). From these sets many maps are calculated, which are averaged and a variance map is obtained. The most intense peaks of this variance map are related to structural variable regions in the biological specimen. The estimation of covariance will help to classify the 3D maps. The other approach is to use maximum likelihood-based classification of 3D models to predict the conformational changes and identify the various molecular states.

1.2.3 Random Conical Tilt Reconstruction

Random conical tilt (RCT) reconstruction method was first developed by Radermacher [13, 14], for reconstructing the macromolecular assemblies without applying symmetry (50S ribosome) [13, 15]. The same field of the sample on carbon support is recorded in pairs with the first image recorded at high tilt (typically between 450 to 600) and the other image at 0° tilt. Due to faster radiation damage and difficult to tilt the sample to high tilt angles in cryo, mostly the tilt pairs analysis is done in negative stain due to its operation at room temperature. Both views of the same area are tracked and aligned with help of recognizable point features. The tilt angle to be used can be roughly decided according to the preferred orientations on the carbon support. In this situation the orientations are more straightforward, since all the tilt views represent the projections with orientations of the cone and untilted views of the particles represent only the in-plane rotation. After classification and alignment of the untilted views of same in-plane orientations and applying the change in two in-plane angles of each view to its corresponding tilted view will determine the position on the cone. This knowledge to determine the orientations of tilted particles is sufficient to reconstruct the first 3D map or group of maps calculated for each class with same in-plane orientation. The missing cone is observed in the data due to limitation in deciding the maximum tilt angle of grid (specimen) holder and also thickness of the tilted specimen. These limitations will affect the resolution mainly in Z direction and can be avoided by aligning and averaging the different conical tilt reconstructions [7]. The resolution of the reconstruction can be improved by estimation of the defocus gradient along the tilt axis and

correcting the loss due to Contract Transfer Function (CTF) [14, 16]. To avoid missing cone, the orthogonal tilt reconstruction (OTR) method has a requirement of well distributed out-of-plane orientations [17]. In OTR method the tilt pairs are collected at $+45°$ and $-45°$ and suppose two particles are apart by out-of-plane orientation of $90°$ with coincident of rotation axes, then $45°$ of one particle will correspond to $+45°$ of other particle. Unlike conical tilt, this method doesnot requires well distributed out-of-plane orientations. This method fills the missing cone and improves the resolution in Z. OTR and its application to a biological sample is shown in [17].

1.3 Algorithmic procedure for single particle reconstruction

The algorithmic procedure for single particle reconstruction can be summarized as follows

1. particle picking: select particles from the images that vary in size and box them manually or automatically;
2. class averaging: The boxed out images are classified and the similar view images are aligned and averaged to obtain better signal to noise ratio;
3. orientation determination: Initial model building or an ab initio estimation are obtained using the random conical tilt technique [18], or common-lines based approaches [19];
4. 3D reconstruction: 3D tomographic inversion of the determined orientation images would generate a 3D volume. This volume is used for iterative refinement in the next step;
5. iterative refinement: The 3D volume reconstructed from the last iteration is reprojected and each raw image is matched to calculate the pose parameters for a new 3D structure and this process iterated until convergence.

1.4 High resolution structure of helical polymers

Helical filaments were the first structures to be reconstructed in three dimensions from electron microscopic images. Helical polymers are made up of proteins, which belong to bacterial, archaeal and eukaryotic cells [7]. Plenty of proteins exist as helical polymers,

which are a consequence of simplest rule of bonding that can be created between two asymmetric protein units [20]. Repeating this bonding rule by many times generates a helical lattice. To reconstruct any protein to near atomic resolution by SPR procedure or 2-D crystal processing requires a lot of effort and one of the important requirements is to avoid missing cone or obtaining all the views of the protein. In this sense, the helical polymers are ideal objects for EM study, because a single image provides all the views required for the reconstruction of the polymer in 3D. The simplicity of the reconstruction procedure of helical polymers in 3D is the reason why the helical polymers have been the first 3D structure reconstruction from EM images [21]. For more than three decades most of the helical reconstructions are involved with Fourier-Bessel (F-B) inversion method [22, 23]. In this method the helical object in Fourier space is described in terms of Bessel functions. The 3D Fourier transform of a helical polymer is nonzero only on discrete layer lines due to the axial periodicity in real space. F-B inversion method requires indexing the layer lines or diffraction pattern by assigning an order n to each Bessel function and then do F-B inversion to real space [24]. Even though the simplicity of the F-B inversion method, there are some serious limitations to this procedure. The alternative is single particle base helical reconstruction method; it became apparent in the past 10 years due to its many advantages compared F-B inversion method [25–27]. It avoids the problem to assign more than one Bessel function on an overlapped Bessel layers [28], it eliminates the straightening of flexible polymers [29], it is able to resolve weakly diffracting filaments [30] and it can deal with the structures that does not maintain long-range order [31]. Also, many different strategies can be applied to even heterogeneous datasets. Invention of Direct Electron Detectors made it possible to eliminate the intrinsic ambiguities of helical filaments and due to all the advantages mentioned above, it is now possible to solve the structure to near atomic resolution.

1.5 Structure and Aim of Dissertation

1.5.1 High-resolution structure of PYD filaments

The aim of this project is to determine the structure of PYD Filaments to the best possible resolution to complement the NMR data to reach atomic resolution. The PYD protein filaments are very thin, so sample preparation should be optimized to go to high resolution in TEM. The aim of the project is to go to resolution $7\sim$ 3Å and use the NMR spectroscopy data to determine atomic structure details for answering

biological questions. PYRIN domain is a protein module that likely mediates protein-protein interactions in apoptotic and inflammatory signaling pathways. Its fibrillation during the apoptosis (cell death) is not well known structurally. The Solid-state NMR spectroscopy data reveals that fibrillation of PYRIN domain is well structured. Due to lack of advanced processing techniques and complexity in structure determination in Solid-state NMR spectroscopy field, the high-resolution electron density map is used to obtain NMR structure with the help of software packages like MOSAICS-EM. The manuscript with this results are included in this thesis (Chapter 5) and the methods related to helical image processing is explained in Chapter 3.

1.5.2 EM Analysis of the basic units of the SIR complex

Our very first goal was to visualize the SIR complex in 3D by transmission electron microscopy (TEM). However, we had to realize that the assemblies formed were of high complexity. Therefore, we focused on the "basic" building blocks of the SIR complex: full length Sir2-Sir4 dimer and tetramer and Sir3 dimers. We applied automated image acquisition with help of Leginon software [32] adapted to our T12 microscope and automated particle picking to average over 20000 particles per complex. For all three complexes analyzed we obtained a range of averages, indicating that the complexes formed are flexible, highly heterogeneous. This also prohibited validating the 3D structure of the complex. For this project, EMAN2 [33]-Single Particle image processing is used to determine the class averages of Sir Complexes. In this technique many separate images are taken, with different views of the same protein and they are aligned to get averages. By Iterative MSA based reference-free and K-means classification the data are classified according to different confirmations and to also reveal the heterogeneity of the sample. The MSA approach is explained in detail in method I (Chapter 2) and the results and discussion are included in Chapter 4. The applications of this automated pipeline for quantitative electron microscopy is presented in Chapter 6-7.

Chapter 2

Method I - Visualize the Heterogeneity by 3D Electron Microscopy

2.1 Abstract

Structural knowledge of many important protein complexes is still limited because of its flexibility, large sizes, complicated architecture etc., Only the structure of the stable and homogenous protein complexes have been solved to high resolution by Electron Microscopy. The methods used for determining the 3D reconstruction from electron microscope images of these types of complexes must consider the heterogeneity of the particles. The method explained here is proven to be feasible to adapt for the study of many more heterogeneous preparations and are appropriate for the 3D reconstruction of macromolecular assemblies from electron microscopic images when ever the heterogeneity may be present.

2.2 Introduction

In past few years, various advancements in technical and computational methods have contributed to the different processing steps in single particle reconstruction (SPR). Several automated particle selection techniques have been developed to pick hundreds of thousands of macromolecules from the images and in some methods the particles can be picked automatically [34]. Advanced particle selection algorithms are implemented based on template matching, neural networks etc., and these techniques are reviewed in

detail [35]. Multivariate statistical analysis (MSA) [35, 36] with multi-reference alignment (MRA) and iterative reference-free alignment [35] with K-means clustering [37] are the two popular methods used to generate class averages. In the process of 3D reconstruction, the determination of orientation parameters is easy in random conical tilt (RCT) method, which in then suffers from missing information due to restricted range orientations. Other than RCT, the Common Lines method [19] is used to build an initial model and determine all orientations required for 3D reconstruction by its projection images. But accuracy of detecting common lines is very low due to low SNR of the EM images. In order to improve the SNR the nearby views are recognized, aligned and averaged to detect the common lines. Once the initial model is determined, it is iteratively refined [38] to obtain better resolution. In every iteration the initial model is updated with the refined model obtained from the previous iteration, this 3D model is re-projected for estimation of the pose parameters to produce a refined 3D model and this process is iterated until convergence.

Independently of the source of the heterogeneity, the variations are carefully analyzed and the particles are classified according to same conformation for calculating a consistent 3D structure. To achieve this task there are three classification schemes, which are sophisticated. Random Conical Tilt (RCT) reconstruction [14, 39, 40], Orthogonal Tilt Reconstruction (OTR) [18] and Electron Tomography [41–44] are the three techniques to generate 3D structures of the different conformations from a single data set. In the process of the Single Particle Reconstruction the RCT method is the most appropriate technique to use and fits well into the pipeline of the entire image processing packages. The 3D reconstruction techniques like the popular Angular Reconstitution method are based on the alignment of projections to an initial model. But the performance of these techniques is doubtful for heterogeneous datasets even if the method uses a 3D reference structure. One of the essential requirements of 3D reconstruction process is to generate a homogenous subgroups by classification or sorting of the particle images. Principle Component Analysis (PCA), Correspondence analysis [45] and Self-Organizing maps [46, 47] are commonly used for Pattern analysis and it is often followed by Classification techniques. All these classification methods are heuristic based approaches and the images are sorted according to the minimization of a mathematical principle and it is minimized. Still these methods are not optimal for sorting a specific set of images. There is still need to carefully supervise any image classification and even apply several classification strategies to get convincing results.

2.3 Class Averaging Techniques

Each projection I has a corresponding 3×3 unknown rotation matrix R which describes its orientation. Considering the ideal case without the contribution of noise, the pixel at (x, y) has intensity $I(x, y)$ which is calculated by line integral of the electric potential induced by a molecule along the path of imaging electrons which is represented mathematically as,

$$I(x, y) = \int_{-\infty}^{+\infty} d\phi(xR^1 + yR^2 + zR^3)dz \,, \tag{2.1}$$

here $\phi : \mathbb{R}^3 \to \mathbb{R}$ is the electric potential of the molecule in some fixed "labora- tory" coordinate system. The projection operator 2.1 is also known as the X-ray transform [37, 48]. The third column R^3 of the matrix R is related to the directions of image, which is called as viewing angle of the molecule. We therefore identify the third column R as the imaging direction, also known as the viewing angle of the molecule. We will often refer to the viewing angle of R as v: that is, $v : SO(3) \to \mathbb{R}^3$ is given by $v = v(R) = R^3$. The viewing angle v can be realized as a point on $S2$ (the unit sphere in $\mathbb{R}^3$) and can therefore be described using two parameters. R^1 and R^2 are the first two columns of the orthonormal basis on the plane in R^3 but perpendicular to the viewing angle v. When some in-plane rotation is operated on the clean projection images of the molecule that share the same viewing angle v, then all those images will look the same. If R_i and R_j are representing two in-plane rotations with the same viewing angle $v(R_i) = v(R_j)$, then R_i^1, R_i^2 and R_j^1,R_j^2 also represent two rotations which represent the same viewing angle $v(R_i) = v(R_j)$. So, R_i^1, R_i^2 and R_j^1, R_j^2 are two orthonormal bases for the same plane and the rotation matrix $R_i^{-1}R_j$ is represented as

$$R_i^{-1}R_j = \begin{pmatrix} \cos\alpha_{ij} & -\sin\alpha_{ij} & 0 \\ \sin\alpha_{ij} & \cos\alpha_{ij} & 0 \\ 0 & 0 & 1 \end{pmatrix} . \tag{2.2}$$

The main aim of class averaging is to increase the signal to noise ratio of the averagely aligned noisy images within a classified group or cluster. These clusters are termed as "classes". In each class all the images should have similar viewing angles, which may also include mirrored images obtained sometimes by artificially double it. These class averages have higher signal to noise ration and similar approach is also applied for later stages of SPR procedure such as angular reconstitution procedure [49]. The quality of the reconstruction depends on the consistent of the class averages and finding it is complicated due to high level of noisy raw images. The rotationally invariant K-means clustering procedure [19] is one of the methods to identify images that have similar

views or viewing angles. The Euclidean distance between the images after alignment with respect to in-plane rotations, which is defined by the invariant distance d_{ij} between the images I_i and I_j and is represented mathematically as

$$d_{ij} = \min_{\alpha \in [0,2\pi)} ||I_i - R(\alpha)\, I_j|| \,, \tag{2.3}$$

where $R(\alpha)$ is the rotation operator of an image by an angle α in counterclockwise direction. Before computing the above mentioned invariant distance, the centering of all images by correlation with template created by averaging the total images as $\frac{1}{n}\sum_{i=1}^{n} I_i$, which looks like radial due to the random rotations of the images. After this step the centers of the images usually off by true centers by few pixel. So, the challenging rotational alignment problem is mentioned in [37] and more review about class averaging methods in [50].

The 2D averages of the images are a fast way of improving the quality of the data and it is calculated by alignment using cross-correlation [51, 52]. It is hard to come up with good quality references for class averaging. It is also important to be careful here, since aligning the set of images by cross correlation with the reference image will bias the data set towards its properties [53–55]. To avoid this problem 'reference-free' alignment is the most commonly used procedure to get the various views presented in the EM data. These procedures by classification are implemented in [56] and most cited method is implemented in SPIDER [57, 58], which includes iterative reference-free technique. The reference-free class averaging is explained in detail in section 2.3.1. For a given roughly centered images, the number of classes is related to the average number of images per class and this parameter can be tuned by observing the final class average of all classes.

2.3.1 Iterative reference alignment

MSA is used to squeeze and remove the noise of the images with the help of 24 – 69 Eigen images. The mathematical explanation behind the MSA procedures is explained in [59]. This algorithm is included in many software's like EMAN2 [33] and IMAGIC [40]. Combining the moving elements post-processor with hierarchical ascendant classification after the data compression is included in SPIDER software package [57, 58] . The two main objectives of this algorithm are to minimize the internal variance between the images in each class ("Intra-class variance") and maximize the variance between the means of the classes ("Inter-class variance"), both simultaneously during each iteration.

In all the steps of single particle reconstruction, the 2D classification and averaging step is very important. Invariant features for initial classification are used for 2D class averaging in EMAN2 [33]. This is a two-step process, in the first step the self-correlation

function (SCF) [60] is used to convert it into translational invariant and then the polar transformation and finally calculating 1-D autocorrelations to each ring for producing SCF images, which are rotationally invariant. Since these images are invariants, it is helpful to bootstrap the process and finally, MSA/MSA based 2D classification gives final result. This process is iterated completely until a stabilized classification is approached. In detail step-by-step procedure of iterative MSA-based reference-free classification algorithm processes as follows

1. initialize the iterative process by making some initial guesses at class-averages. These are invariant-based, meaning that even with MSA, this initial classification is not exceptionally good.
   (a) For each particle, rotational/translational invariants are calculated;
   (b) apply MSA on the invariants to define an orthogonal subspace, which reveals the most important features (differences) among the classes;
   (c) these particles are re-projected into the subspace (MSA) using basis vectors and let say "nbasis";
   (d) K-means based Classification and it is classified into used defined number of classes i.e. "ncls";
   (e) initial averages are generated after an Iterative class averaging of the particles;
2. these class averages (current) are aligned to each other, sorted and also kept centered;
3. now MSA is performed on the aligned class averages images instead of invariant and it will represents largest variations;
4. from the above class averages a subset is selected according to used defined "naliref" as an alignment references for this iteration;
5. in this step each particle is aligned to each reference (averages from the last step) and the orientation of the particle related to best-correlated reference is stored in database;

6. project aligned particles using reference MSA vectors from step 3;

7. the projections produced from step 6 are applied to K-means classification;

8. new class averages are generated according to groups obtained from previous step and it is used for next iteration;

9. go to step 2 and iterate until user defined number of loops "iter" are completed.

2.3.2 Contrast transfer function estimation and correction

The images obtained from electron microscope are phase contrast images and considered as good approximation of the same. These images are recorded with under-focus or defocus (the distance of the object from the focal plane) and combined with the spherical aberration the required phase contrast is created for visualizing the sample. An electron microscope image is best represented as the convolution of an ideal image with point spread function of the microscope. The Fourier transform of the image is multiplication of the Fourier transform of the perfect image with the Fourier transform of the phase-contrast transfer function (CTF). Depending on the periodicity and amplitude of the CTF, it enhances and suppresses different features of the object [46]. The CTF function can be described by the equation

$$H(k) = \sin\gamma(k) - W\cos\gamma(k)\,, \tag{2.4}$$

were W is the amplitude contrast ratio, and $\gamma(k)$ is the phase shift produce by the lens aberration and defocusing. The amplitude of oscillations of CTF varies between -1 and 1 as a function of spatial frequency (Figure 2.1A. The *envelope function* has to be considered during CTF corrections in image processing due to its frequency-limiting impact on the EM image (Figure 2.1B. The CTF function will have constant information until the first zero crossing which normally is low-resolution regime and this is dependent on the defocus value of the image. At high resolution, the CTF will oscillate from +ve to –ve introducing a lot of zeros and it is essential to correct for these phase reversals to retrieve the high-resolution information. To fill in the zero information caused by oscillations the images of the same object are recorded at different defocus, so that these zeros are different between the images. Even before considering fitting CTF parameters for individual images, images should be assessed qualitatively for drift, astigmatism, and overall resolution. Studying the 1D and 2D power spectrum in detail reveals the quality of the parameters explained above.

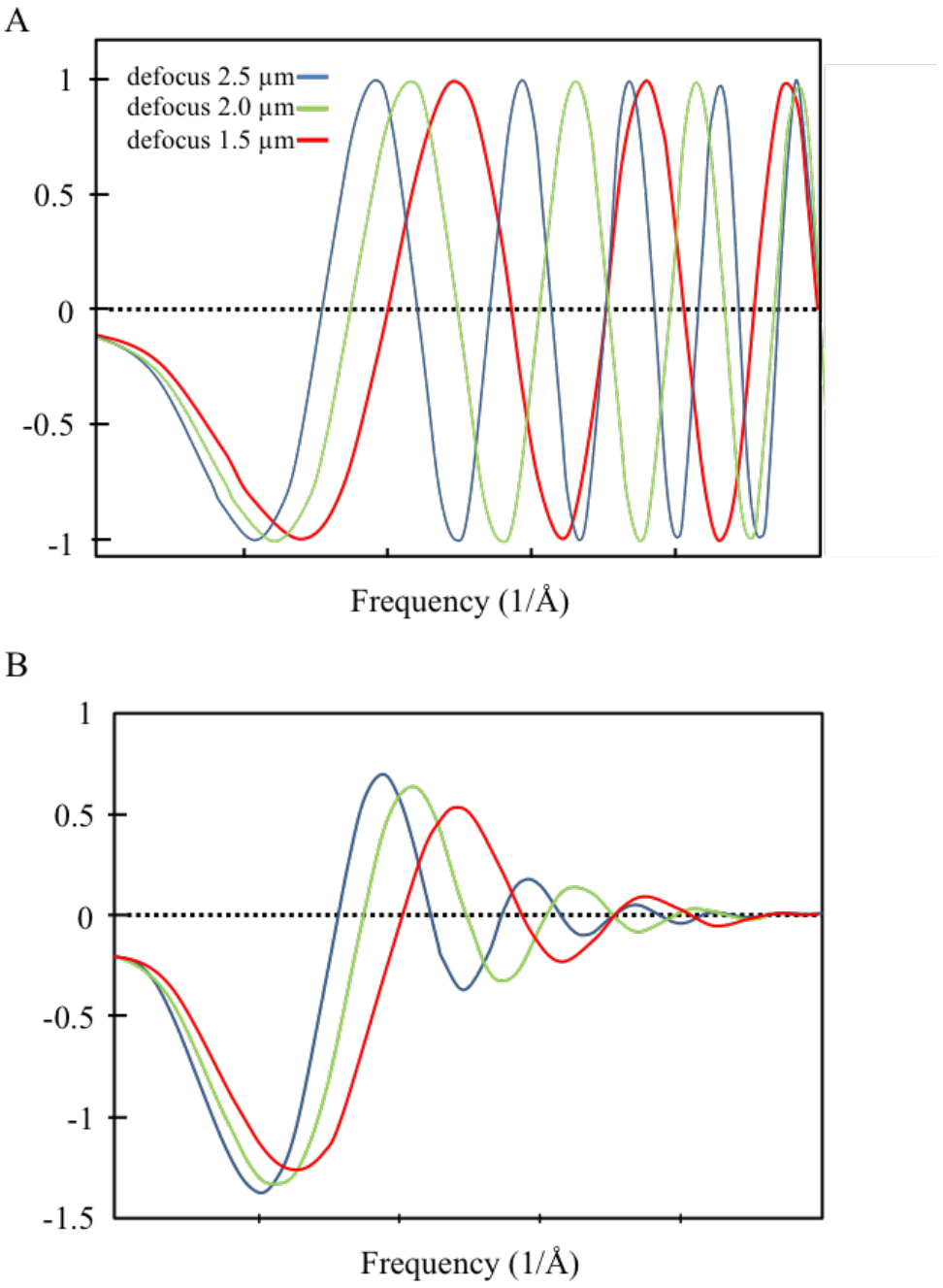


FIGURE 2.1: The contrast transfer function (CTF). (A) A theoretical CTF for 3 defocus values: $1.5\mu m$ (red), $2.0\mu m$ (green) and $2.5\mu m$ (blue). The oscillations of this function vary from -1 to +1 and the position of first zero crossing is considered as defocus value. For lower values the information loss is minor compared to higher values of defocus and to avoid this several images of identical structures are taken at different defocus values to cover all the frequencies and "fill in" in the information lost by zero crossings. (B) A theoretical CTF with envelope function effects, which has to be considered during image processing.

2.4 Random Conical Reconstruction

For any new sample of unknown structure a reliable initial 3D model is necessary and continue the refinement with Projection matching. Nonetheless Random Conical Tilt resonstruction (RCT) [18, 57, 61, 62] technique is well established and there is no need for prior knowledge about the structure and once the 3D structure is calculated the projection matching is used to refine it to high resolution. RCT technique applied to a heterogeneous sample for classifying it into different 3D structures to understand its complex behavior or mixture.
Principle of the Random Conical Tilt method:

1. The principle of RCT method requires a data set consisting of pairs of micrographs of the same area of the sample on a carbon support with first image recorded at higher tilted angle, typically 55° to 65° and second image recorded at 0° (Figure 2.2 A - B).

2. Pairs of the same particle are picked from the micrographs by keeping it side by side (Figure 2.2 C).

3. The CTF estimation and correction is performed on both tilted and untilted images as explained in Section 2.3.2.

4. Every particle in the tilted image provides a other view of the same particle in untilted image, but each from a different direction. In 3D each pair is correlated by the direction of the tilt axis and its tilt angle, which form a conical tilt series (Figure 2.2 D).

5. The untilted particle images are aligned and classified into different homogeneous groups. Each aligned untilted particle brings the matching tilted particle image to its corresponding azimuthal location (Figure 2.2 E).

6. The 3D structure of the molecule is reconstructed using back-projection algorithm with all the tilted particles in each group of classification, filling the azimuthal space. For heterogeneous data sets the 3D reconstructions of same classes with different orientations are merged and those with different structural features are left as separate reconstructions for further analysis (Figure 2.2 F).

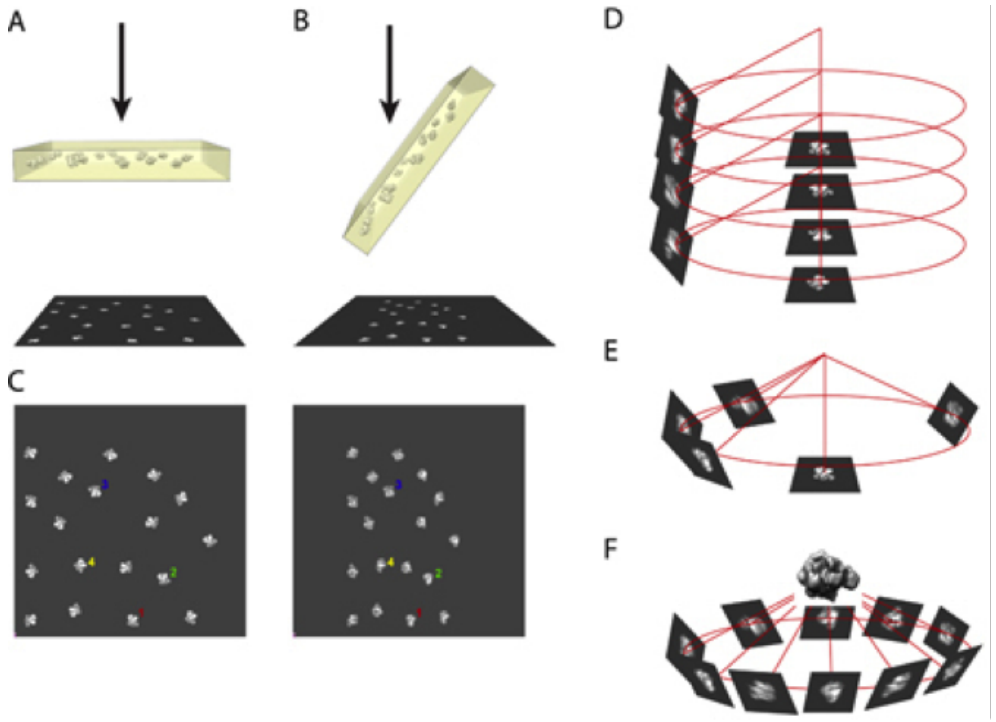


FIGURE 2.2: Illustration of RCT technique [63]

2.5 Tilt pair particle picking

The particles in the tilt pairs are efficiently picked using programs in EMAN2 [33] package. This package gives flexibility to correct false particles picked in the pair of images interactively and semi automatically with a convenient GUI tool shown in Figure 4.10. The user needs to pick first five matching pairs of particles and from sixth particle onward the user needs to pick the particle only on the untilted images. This works by calculating the geometric relation of both images from first five matching pairs. The computation of the geometric relation between the tilt pairs is a two-step procedure. First, the tilt angle is calculated by the difference in the area of the triangles defined any three particle coordinates. The relation between the area X0 of the triangle in the untilted image and the area Xt of the triangle in the tilted image given by

$$X_t = X_0 \cos(\theta) \,, \tag{2.5}$$

where θ is the tilt angle. From a 'n' picked particles there an N triangles that can be calculated

$$N = \frac{n!}{3!\,(n-3)!} \tag{2.6}$$

A parameter to eliminate smaller area triangles is used which avoids inaccuracies in particle picking and also in the calculated tilt angles. Using Least squares fit the rotation

and translational parameters are calculated by fixing θ. At the end we have 10000 measurements to calculate the tilt angle at an accuracy of less than 0.10.

2.6 Processing of untitled images

The processing of 0° images can be done with any SPR processing package with for reliable 2-D image alignment and good classification techniques, in our project for SIR complex we used all the tools available in EMAN 2.0 package [33]. Any package should take care of not to corrupt the information of correspondence between the tilt pairs with some naming convention. This correspondence should be maintained even if subsets of images are removed by following the naming convention. If the tools used in packages do not support this, then 3D reconstruction from RCT or orthogonal tilt reconstruction is complicated.

After particle picking from tilt pair images, the particles from 0° tilt images are processed first and this step is critical part for analyzing the heterogeneity of the data set by random conical tilt. In this step the particles are aligned and classified into different groups that have identical particle. The reliable classification of particles depends on the classification technique used and so, for very noisy data set aligning image to a common reference may lead to bias to the alignment. Therefore, a reference free alignment method as described in section 2.3.1 is used. The number of classes to be generated depends on the amount of data set and for heterogeneous sets a trade of between the number of classes and the number of particles in each class is necessary to maintain good SNR of the class average. This decision can be easy by keeping in memory the structural variations in the sample at all the steps from data acquisition to alignment and classification. Finally, after processing the 0° particles we can get the information about in-plane rotation angle of each particle, including the class number. This information is transferred to the set of respective tilt particles as azimuthal angle in the conical tilt series. During the particle picking process the fixed tilt angle between the tilt pairs are stored which need to be close to readout during the data collection in the microscope. This leads to the ability to calculate the 3D reconstruction from each class separately.

2.7 3-D reconstruction by Radon inversion

Radon inversion is a linear algorithm [64–66] highly accurately and fast enough to calculate the 3-D reconstruction from randomly oriented projections [66]. For RCT reconstruction, the Radom transform provides many advantages due to its applicability of the

central section theorem [67]. Normally, central section theorem applies to the Fourier space, but the radon transform has the advantage of using real space to reduce the complexity of computation by using simple interpolation techniques. A precise interpolation in Fourier space needs kernels that can cover large areas of the transform, such as moving window Shannon interpolation [68] A simple linear interpolation, if not even nearest neighbor interpolation is enough to interpolate Radon transforms. Because the convertibility to Fourier space at any point of time for calculating cross-correlations, it is perfectly feasible due to simple calculation of 1-D Fourier transform along the lines of multidimensional Radon transform which results in a multidimensional Fourier transform in a polar coordinate system.

Chapter 3

Method II - Processing Images of Helical Structures

3.1 Introduction

An enormous variety of regular structures is found in nature. As shown in (Figure 3.1) regular, three-dimensional (3D) structures that are assembled from a single type of subunit have one of three forms. Each form is most conveniently described in the

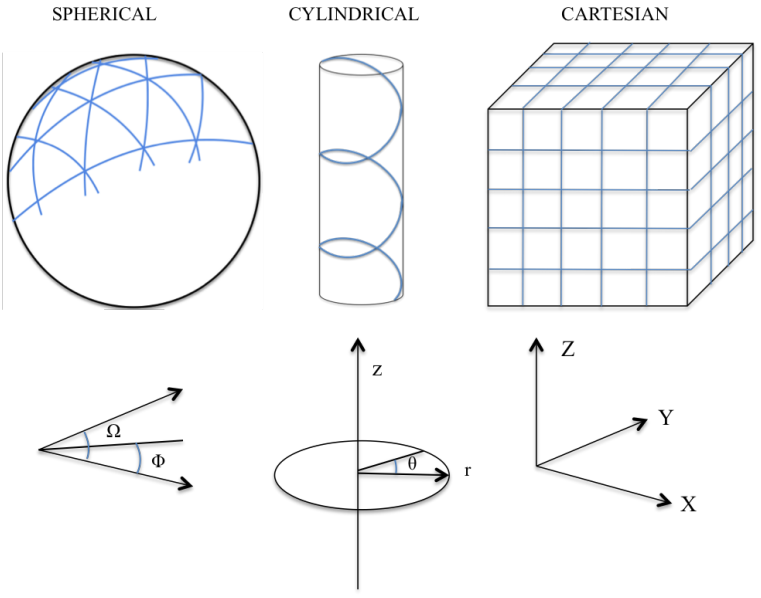


FIGURE 3.1: The three basic types of regular structure together with the respective coordinate systems.(*courtesy:Prof.Dr.Andreas Engel*)

corresponding coordinate system. Cylindrical coordinates are appropriate for helical

structures. In this case polar coordinates (r,θ) are used in the plane x,y perpendicular to the helix axis,z. The most well known helical structure in biology is the DNA double helix. To solve its structure, an event that represents the foundation of modern biology,the nature of helical diffraction had to be understood. Because the subunits of a helix are not only axially but also angularly displaced with respect to each other, a single projection of a helical structure reveals various projections of the subunit, making it unnecessary to collect tilt series (as for sheet structures) to achieve a 3D map.

3.2 Real Space reconstruction of helical filaments

IHRSR is based on the fruitful single particle approach to determine macromolecular structure, which are asymmetric or low symmetric structures. It retains the same strategy as single particle approach but modifying it, so that it can refine helical symmetry parameters during the iterative procedures for structure determination.

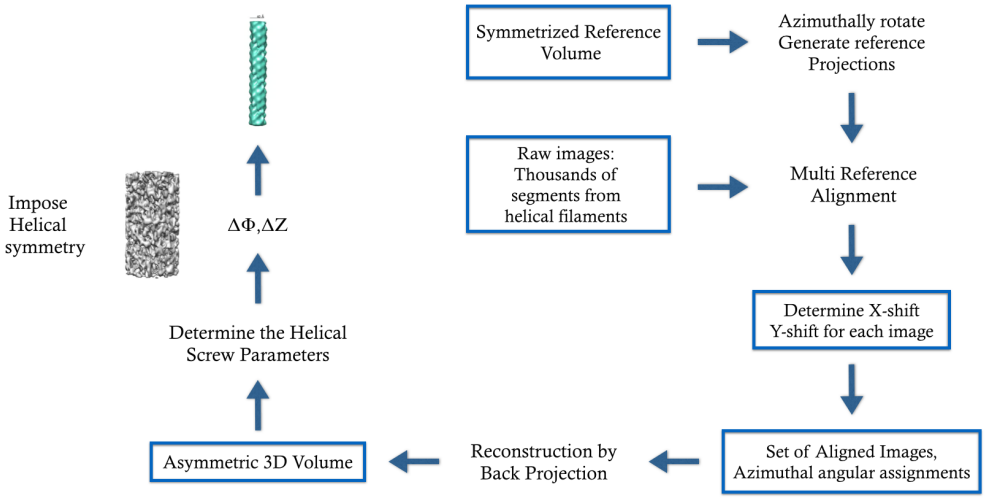


FIGURE 3.2: Iterative Helical Real Space Reconstruction

Requirements:

1. Good initial guess of the helical symmetry parameters.
2. Initial 3D template (noisy cylinder).

It can deal with

1. Heterogeneous structures.
2. Weakly diffracting filaments.

This algorithm is enormously robust due to high degree of averaging and regularization properties of helical symmetry imposed. The procedure (Figure 3.1) begins with a symmetric reference volume. The script generates a series of projection of the volume. These projections are aligned with raw data (stacked helix segments). After finding the best fit, the program uses the data to generate an asymmetric 3D volume. Finally, helical symmetry is imposed. This results in a new symmetric volume, which is used as a starting point for the next cycle of IHRSR. Since there is no knowledge of the structure at the beginning, the very first cycle begins with a solid cylinder. Modified from [25].

A new implementation of IHRSR in SPARX [38] has lot of features intended at improving refinement procedures and convergence properties. The scripting capabilities of SPARX and advanced image processing techniques are used to make the decade old original version modernized [69].

The iterative process as explained in 3.2 has to converge by observing:

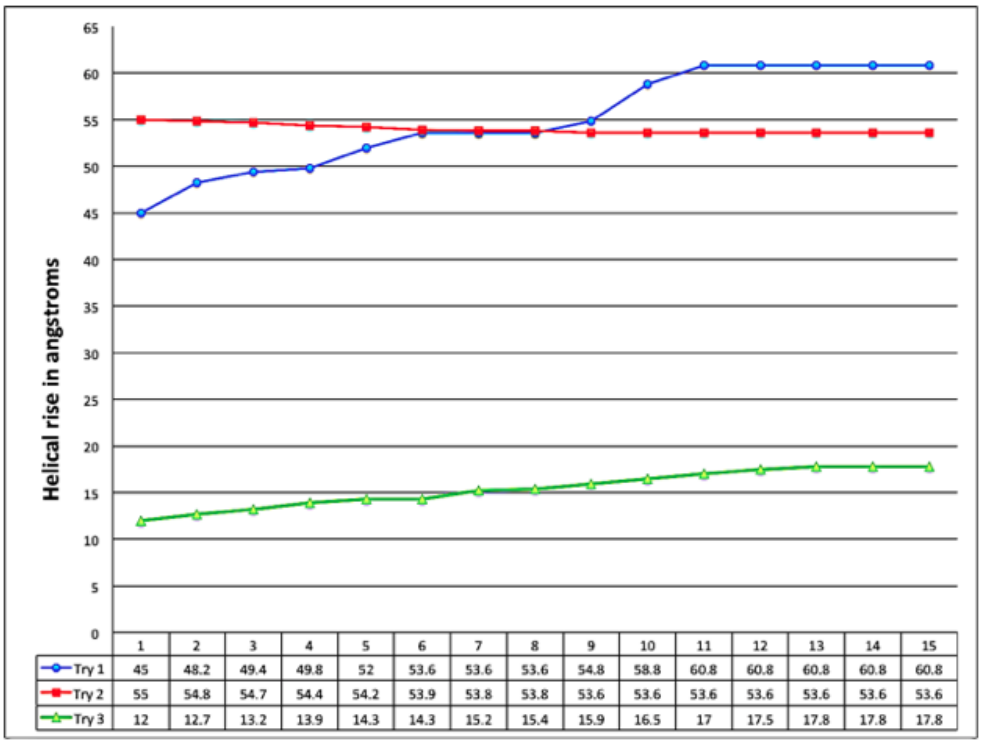
1. The assigned angles for the projection images, which are expected to settle.
2. The azimuthal distribution of projection directions is expected to be uniform in the range $0° < \theta < 180°$.
3. Helical symmetry parameters should stop varying.

Even though the algorithm is robust to solve most of the structures,

1. Initial combination of helical symmetry parameters should not be far from the exact values.
2. Ambiguities in helical symmetry – may correspond to quite diverse combinations of helical symmetry parameters.
3. Even after convergence of parameters the solution may not be optimal. Since the algorithm is implemented with simple downhill minimization. The convergence plot of rise and angle between the subunits applied to PYD filaments [Chapter] is shown in the Figure 3.3 and Table 3.4.

Improvements in SPARX package [69] compared to original version:

1. Both Cubic and Non Cubic volumes can be reconstructed with help of modified reciprocal space interpolation scheme to use only the central rectangular real-space portion of 2D projection image.



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Try 1	45	48.2	49.4	49.8	52	53.6	53.6	53.6	54.8	58.8	60.8	60.8	60.8	60.8	60.8
Try 2	55	54.8	54.7	54.4	54.2	53.9	53.8	53.8	53.6	53.6	53.6	53.6	53.6	53.6	53.6
Try 3	12	12.7	13.2	13.9	14.3	14.3	15.2	15.4	15.9	16.5	17	17.5	17.8	17.8	17.8

FIGURE 3.3: Convergence plot of IHRSR algorithm for helical rise applied to PYD helical filaments (Δz)

$\Delta\phi$	Initial (in degrees)	Final (in degrees)
Run 1	90	90
Run 2	60	60
Run 3	40	40

FIGURE 3.4: Convergence plot of IHRSR algorithm for between sub units $\Delta\phi$ applied to PYD helical filaments with 3-start helical symmetry

2. Proper Wiener filter-like CTF correction is incorporated with reconstruction procedure.

3. Analysis of specimen heterogeneity by Helical PCA.

4. The scripts are parallelized using MPI on a Python level.

3.2.1 Intrinsic Ambiguities in indexing diffraction of helical filaments

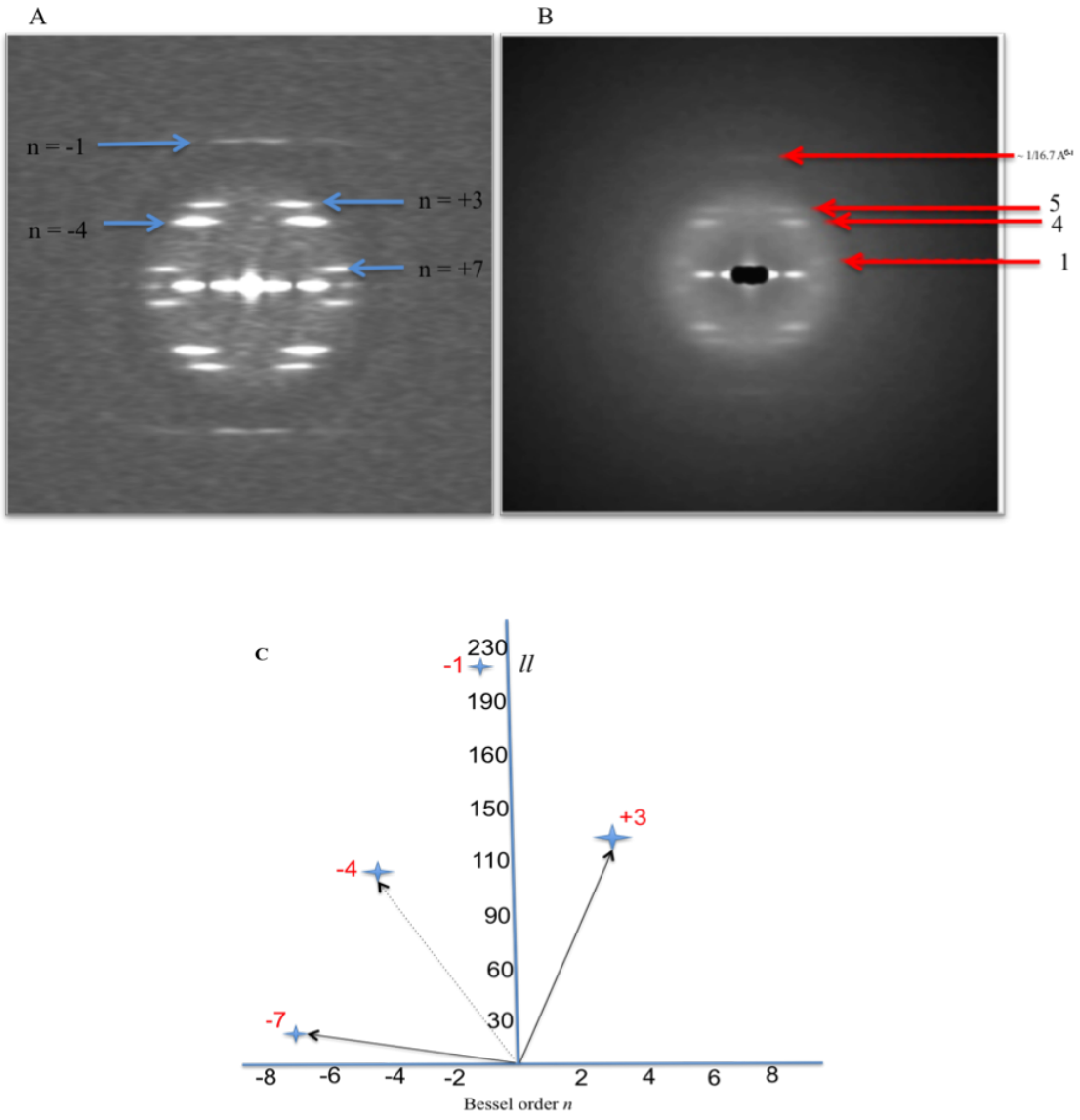


FIGURE 3.5: (A) Average power spectrum of the MAVS filaments from the images recorded using direct electron dector. (B) Average power spectrum of the same filaments from the images recorded in CCD. (C) n,l plot for diffraction pattern of (A)

The performance of IHRSR depends on chosen initial guess symmetry parameters. As an example, the average power spectrum of MAVS-CARD [70] filament images from CCD and Direct Electron Detector are calculated. The possibility of two solutions for MAVS-CARD filament structure is understood better with help of power spectra. The power spectrum in (Figure 3.5A) is the correct solution for MAVS filament which reconstructed at $\sim$ 4Å [70]. The layer lines are indexed exactly as shown in [70]. The average power spectrum of same sample at low resolution predicts that the n=0 layer

line is $1/16.7\text{Å}^{-1}$ (Figure 3.5B) which corresponds to rise between the subunits ($\triangle z$) and when applied results in a wrong 3D reconstruction. This value for rise ($\triangle z$) contradicts the layer line analysis of (Figure 3.5A), which reveals that the $1/16.7\text{Å}^{-1}$ layer line is correspond to "-1" layer line. The layer line of this power spectrum are indexed as shown in n,l plot. Each point in this plot is a "n" start helix. Infinite number of points corresponds to infinite number of helices. The two arbitrary vectors shown in the (Figure 3.5C) are used to plot every point in the n,l plot by linear combination of these two vectors. In [70], it is shown that the first meridional point is at $1/5.13\text{Å}^{-1}$. So, to predict this correctly high resolution information is necessary and the only way to validate initial guess parameters is to refine the structure to less than $\sim$ 6Å to see the alpha helices in the reconstructed volume.

Chapter 4

Sir4 dimerization promotes oligomerization of the SIR complex

Stephanie Kueng†, Venkata Prasad Dandey[1], Mohamed Chami[1], Mariano Oppikofer[1],; Monika Tsai Pflugfelder, Veronique Kalck, Simon Weidmann[3], Renato Zenobi[3], Henning Stahlberg[1] and Susan M. Gasser[2]

Friedrich Miescher Institute for Biomedical Research, Maulbeerstrasse 66, 4058 Basel, Switzerland
1) C-CINA, Mattenstrasse 28, 4058 Basel, Switzerland
2) University of Basel, Faculty of Natural Sciences, Klingelbergstrasse 50-70, 4056 Basel, Switzerland.
3) ETH-Zurich, Switzerland
†current address: Hoffmann-La Roche Ltd. Pharmaceuticals Division, Grenzacherstrasse 124, 4070 Basel, Switzerland

*Corresponding authors: SM Gasser; susan.gasser@fmi.ch
*Corresponding authors: H Stalhberg; henning.stahlberg@unibas.ch

The results presented here are are mainly my contribution and planned to be included in a manuscript, which is in preparation.

4.1 Introduction to SIR complex

Nucleosome core particle (NCP) is the major unit of DNA packaging in eukaryotic chromatin [71]. The octamer of four major histone proteins- H2A, H2B, H3 and H4 wrapped around by DNA forms the NCP [72]. The higher levels of chromatin structure can be built upon the foundation of the NCP [73]. The activation of the specific genes can be facilitated by the organization of chromatin [74, 75]. The covalent histone modifications play a major role in determining the organization of chromatin. One such example is acetylation of histone proteins which results in the weakening of the interaction between histones and DNA resulting in destabilization of the nucleosome structure and thereby, the genetic locus becomes more accessible to the transcription complexes [76, 77]. The repression of the transcription can also be a result of the binding of non-histone proteins [78]. This function is achieved by Silent Information Regulator (SIR) heterotrimeric protein complex, which is composed of SIR2, SIR3 and SIR4 proteins [79]. All three proteins in this trimeric complex are essential for gene silencing. The three heterochromatin-like regions are present in budding yeast viz. rDNA locus; subtelomeric regions and the homothallic mating (HM) type loci- HMR and HML that regulate the identity of the cell [80]. The initial nucleation sites of SIR complex consist of a bridge at HM loci between SIR1 and the transcription factors which are required for recruitment of the SIR proteins at the telomeres and HM loci and also for the recognition of specific DNA motifs at these sites. The complex spreads from these nucleation sites along the nucleosomes to repress nearby promoters [80–82].
The limited concentration of SIR proteins in nucleus and the decreased affinity of SIR3 with chromatin due to histone modifications ceases the spread of the repression of transcription by SIR complex [83–87]. The cascade of events in the silencing starts from the stimulation of the deacetylase activity of SIR2 in vitro by its strong interaction with SIR4. Then, dimerization of C-terminal coiled domain of SIR4 and its binding with SIR3 takes place, which is important for the silencing. SIR3 consists of a 214 (1-214) amino acids long N-terminal bromo-adjacent homology (BAH) domain (Figure 4.1), which interacts with the nucleosome [88–91] and AAA+ ATPase-like (AAA) domain (532-845) with the missing ATPase activity because the residues required for ATP hydrolysis and catalysis are missing in the case of SIR3 [92, 93]. The N-terminal domain (1-214 aa) binds nucleosomes [91, 94], AAA domain interacts SIR4 (mostly 464-728 amio acid residues of SIR3) and the nucleosomes [95], and a C-terminal domain binds to itself and helps in homodimerization (aa 840–978) [96–98].

Chromatin mediated gene silencing in budding yeast is mediated by the SIR complex, which is composed of Sir2, Sir3 and Sir4. This complex is recruited to silent sites by sequence specific DNA-binding partners (e.g. Rap1 or Yku at telomeres), from

where it spreads into nearby chomatin. High affinity binding sites on nucleosomes are generated by the deacetylation reaction of Sir2, which in particular targets H4K16ac, a widespread modification in yeast chromatin. The spreading of the complex is necessary but not sufficient to silence any promoter put in the SIR-bound chromatin region. The phyisological role of gene silencing is to allow sexual life cylce of yeast and the regulation of subtelomeric stress genes. The latter does not seem to be essential under laboratory conditions, but may be important for wild and/or infectious yeast species.

The phyisological role of gene silencing is to allow sexual life cylce of yeast and the regulation of subtelomeric stress genes. The latter does not seem to be essential under laboratory conditions, but may be important for wild and/or infectious yeast species. The Sir2-3-4 complex becomes smaller when Sir3 dimerization is disrupted (Figure 4.2C). Also the Sir4 2xB mutant results in a smaller complex - either because less Sir3 is bound or because Sir4 dimerization is disrupted. These results argue that both Sir3- and Sir4-mediated dimerization occurs.

4.2 Sample preparation

Sir2-4 is purified using Cm-beads; Sir4 is C-terminally tagged with a Cm peptide. This is one-step purification (elution with EGTA) and has been cleaner than HIS or SII tag purifications. A double-purification is tried once (first HIS, then Cm), but the end result was not cleaner, just less concentrated. Since Sir4 is the least expressed component, purification on another tag first reduces the material significantly. Before we started to use GraFIX [99], the proteins were not homogenous at all (gradient cleans up) and the fixation helped to keep it intact. Detergent was included in both, experiments 1 and 2 (Figure 4.3). We can leave it out, but grids from samples prepared without (for Sir2-4 WT) did not improve significantly. Finally buffer composition of material put on grids: 10mM Triethanolamine pH8, 150mM NaCl, 15

4.3 Grid preparation

Initially we tried different kinds of grids and staining: Parlodion/carbon film, floated carbon film. The samples are adsorbed on glow-discharged grids or back injected on freshly floated carbon film. The grids are stained with 2

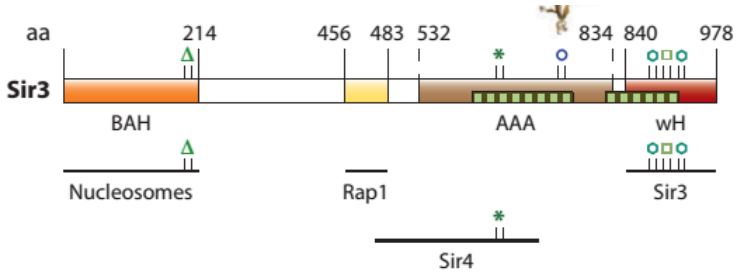


FIGURE 4.1: Color-coded schematic of Sir3 protein with its important domains, protein-protein interactions, and mutations.

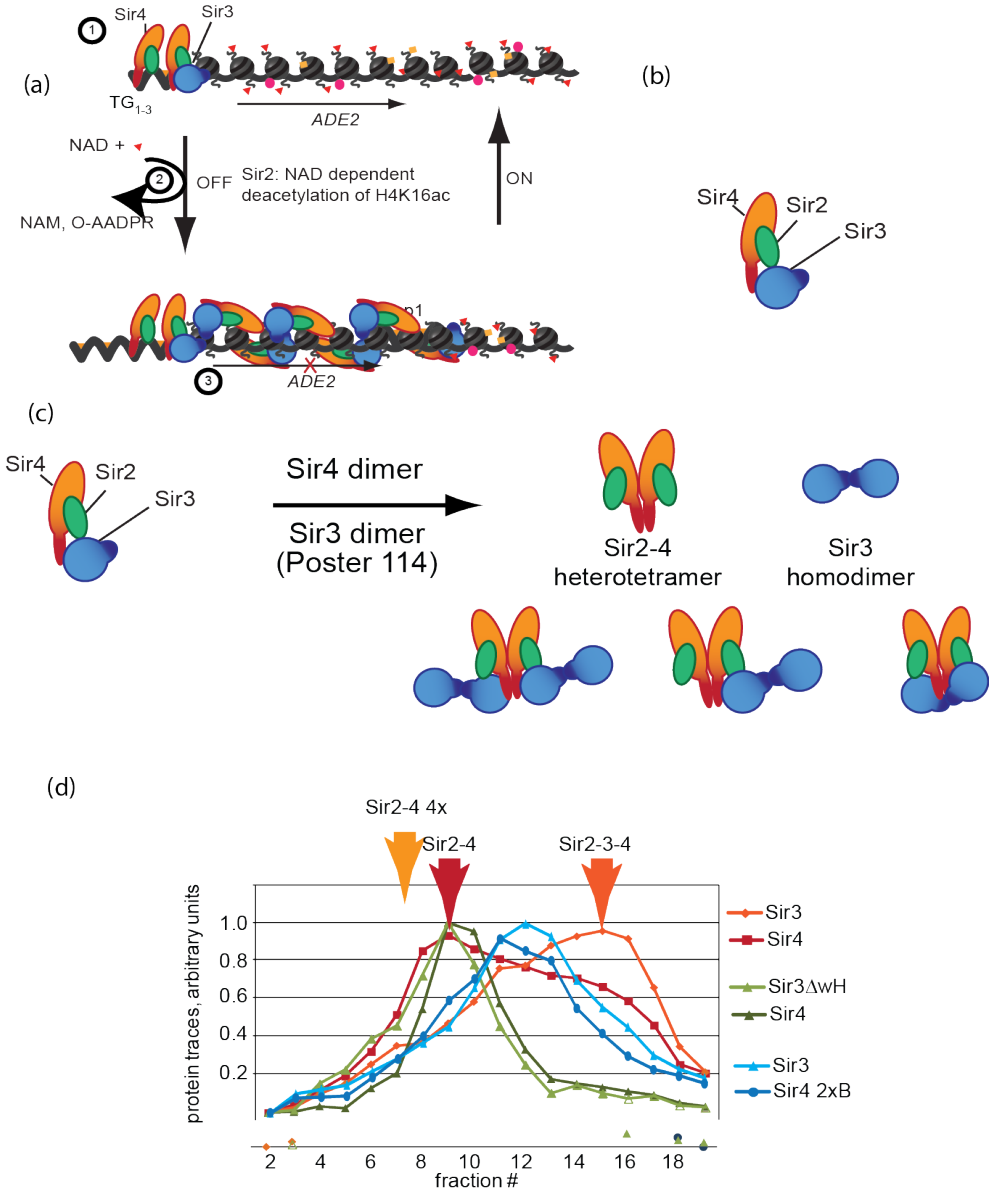


FIGURE 4.2: (a)Schematic diagram explaining gene silencing mediating by SIR complex. (b)Schematic Sir234 complex (c-d)Schematic and results shows the Sir3 and Sir 4 mediation during dimerization.

4.4 MALDI-TOF analysis

In order to determine if we are really looking at a dimer (and not a tetramer or something else), we used MALDI-TOF established by Simon Weidmann in Renato Zenobi's lab at ETH Zürich (Figure 4.4). Out of three fractions, we measured fraction 8 from experiment 2 (it is suspected – but don't know for sure- that the others were too low concentrated). From fraction 8 and conducting an second experiment. The MW of the main peak is 247kDa. Sir2-4 (including tags etc) has 219kDa. The additional MW may come from crosslinking additions. Sir4 alone (without crosslinking) had a MW of ca. 160 kDa, fitting well with the actual weight. Nevertheless, this MW is only explained by Sir2-4 and confirms that most of the fraction contains the dimer.

4.5 Data Collection

4.5.1 Imaging and checking image quality with CM1

Manual TEM was carried out on a Philips CM10 operated at 80 kV. The images were recorded on a 2k x 2k CCD camera (Olympus SIS, Münster, Germany). For 2D class averages, images were acquired at 130000x nominal magnification (pixelsize: 0.37 nm, defocus: 0.1 – 0.3μm. Particles were picked manually using EMAN2 [33] software package, classified and averaged using the e2refine2d algorithm. Before taking automated pictures, the best grids were selected at CM10. These images may also help us to check the quality and origin of the background.

4.5.2 Automated Image acquisition using T12 microscope

Automated image acquisition was done on a FEI T12 operated at 100 kV using the Leginon 2.1 (incl. in Myami 2.1) software, which is explained in(Figure 4.5 [32]. The images were recorded on a Gatan 2k x 2k CCD camera. First, a bunch of squares on the TEM grid were manually selected for analysis. Then, a mesh consisting of 35 sub-squares was created within every square, and 11 images per sub-square were automatically acquired at 54000x magnification. The particle picking was done manually and/or semi automatically using the Appion 2.1 (incl. in Myami 2.1) software [100] and EMAN2 [33] package in combination with a template picking routine [36].

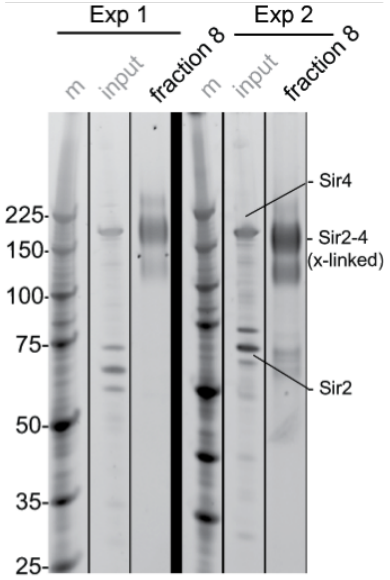


FIGURE 4.3: Co-expression of Sir4-Cm and Sir4-SII (Streptactin II) of either two wild type proteins or two proteins with four point mutations in the coiled-coil domain.

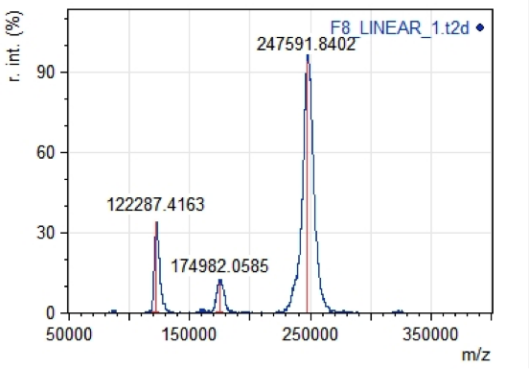


FIGURE 4.4: MALDI-TOF results for Sir2-4 mutant complex shows the sample has mixture of three populations in which Sir 2-4 with 247 kDa by Mol. Wt. is the majority of the fraction.

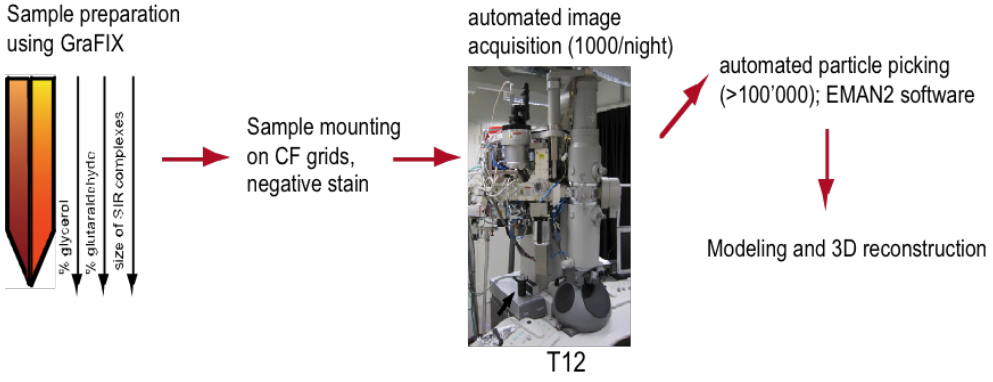


FIGURE 4.5: Semi Automated pipeline adapted to analyse the SIR complex. Stabilizing the sample using Gradient Fixation (GraFIX) method, transfer to Fully automated Tecnai 12 Microscope which records 1000 images/day and automated particle picking to pick 100000 particles followed by Heterogeneity analysis by 2D and 3-D reconstruction

4.6 EM analysis of the basic units of the SIR complex

It has been shown previously that Sir3, similar to Sir4, forms stable dimers in solution [97]. We thus sought to structurally analyze these basic units of the SIR complex, hoping this would allow us to talk the structure of the holo-complex. In order to stabilize the complexes for EM analysis, we used the GraFIX method by which the complex gets slowly crosslinked by EM grade glutaraldehyde while separating different size molecules in a glycerol gradient [101]. This allowed us to purify homogenous complexes that showed quite good contrast by negative stain. We could analyze the Sir2-Sir4 tetramer, the Sir2-Sir4 dimer and the Sir3 dimer. The Sir3 monomer was most likely too small to obtain useful images. There was again a clear size difference visible for Sir2-Sir4 tetramer complexes and Sir2-Sir4-AB mutant forming dimers (Figure 4.6,Figure 4.7, Figure 4.8). In contrast to the non-crosslinked samples the crosslinked complexes did run roughly at their expected size in the gradient, indicating that crosslinking may reduced some of the flexibility. Nevertheless, the complexes did retain various shapes: we applied automated image acquisition and particle picking to average ¿ 4000 particles to obtain more structural information. The resulting averages had various shapes (Figure 4.9), so that we could not obtain a clear 3D volume. Nevertheless, the averages show that the Sir2-Sir4 tetramer (Figure 4.6) is composed of two molecules resembling the Sir2-Sir4 dimer (Figure 4.7) but in different conformations to each other. The Sir2-Sir4 dimer averages have a diameter of 9.5-15 nm, whereas the Sir3 dimers are between 7.5 and 9nm large. The particle size and shape for Sir3 dimers is in agreement with previous reports [101]. We however did not see higher order oligomerization states; the dimer peak was the most abundant in the gradients of purified Sir3 [31]. In [34], the authors did not obtain images for Sir2-Sir4 alone so we cannot directly compare their Sir2-Sir4 + Sir3 preparation to our samples. We have tried to purify the holo-complex applying the stabilizing and size-separating GraFIX method. We were however unable to obtain homogenous enough populations for single particle analysis (data not shown). This most likely reflects the multiple ways of complex formation in combination (see below) with the already flexible basic units reported here.

4.7 Random Conical Tilt Reconstruction with Tilt pairs Sir 24 mutant complex (For heterogeneity analysis)

4.7.1 Data collection of Tilt pairs

The tilt pairs were collected in FEI CM200 microscope equipped with FEG filament with a TWIN-lens pole-piece configuration setup [102] operated at 200 KV and nominal

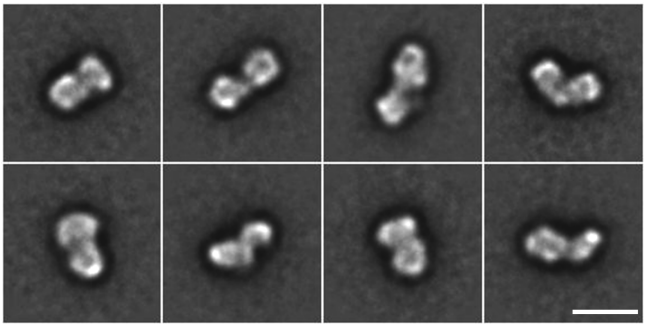


FIGURE 4.6: TEM analysis of negatively stained Sir24WT sample. 4679 images of elongated particles were classified into 18 classes. Classes with fewer than 180 particles or those where an average particle didn't show a clearly delineated border were rejected. From the remaining 12 class averages, 8 representative averages are shown here. Scale bar: 20 nm.

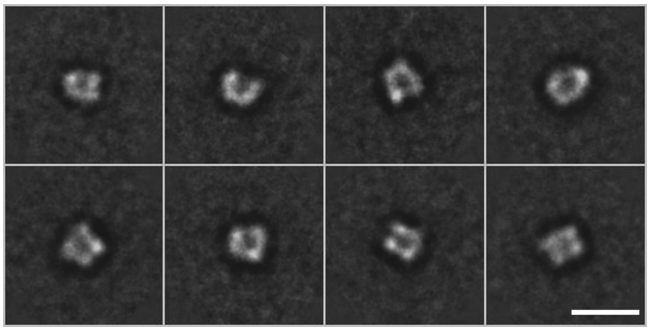


FIGURE 4.7: TEM analysis of negatively stained ir24cc sample.4679 images of elongated particles were classified into 18 classes. Classes with fewer than 180 particles or those where an average particle didn't show a clearly delineated border were rejected. From the remaining 12 class averages, 8 representative averages are shown here. Scale bar: 20 nm.

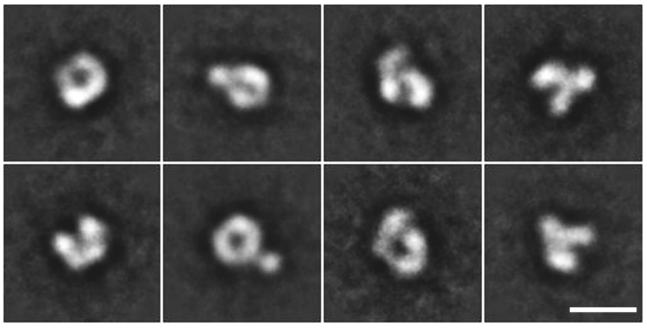


FIGURE 4.8: EM analysis of negatively stained Sir3WT. 4679 particle images were classified into 32 classes (See also Supplementary Data). Classes with fewer than 75 particles or those where an average particle didn't show a clearly delineated border were rejected. From the remaining 24 class averages, 8 representative averages are shown here. Scale bar: 12 nm.

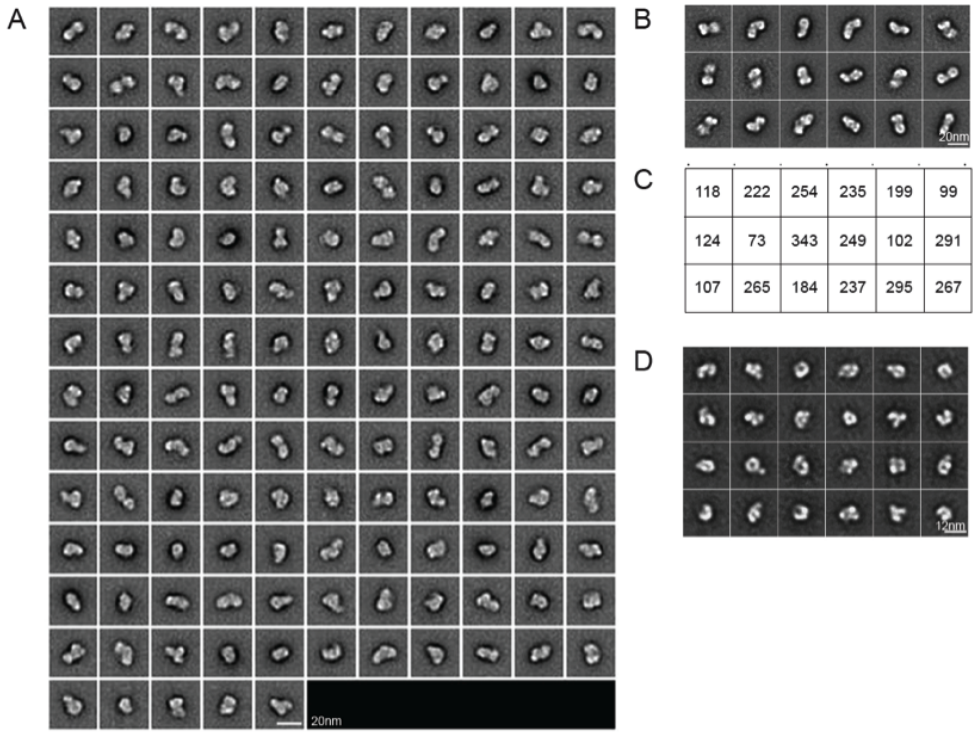


FIGURE 4.9: Full averages of EM analysis of Sir2-Sir4 hetero dimer, hetero tetramers and of Sir3 dimers. (A) Sir2-Sir4 analysis by negative stain TEM. A Total of 22008 particles were picked from 483 micrographs and 200 class averages were generated using EMAN2 [33]. Class averages were rejected for classes with less than 75 particles, which did not have a clearly delineated contour, or that showed faded regions within particle areas. The remaining 128 class averages are shown here. The number of particles in each class was between 90 and 130 particles. Scale bar: 20nm. (B) Sir24 WT analysis by negative stain TEM. A total of 4679 Sir2-Sir4 particles were classified into 18 classes, and class averages were evaluated to remove outliers using e2evalparticles.py from the EMAN2 [33] package. Class averages showing an elongated feature are represented here. A Cut off of 0.80 for class keep (based on similarity score) was used in each class average to exclude bad particles during refinement. Scale bar: 20nm. (C) The number of corresponding particles in each class average for (B). (D) Sir3 dimer analysis by negative stain TEM. A Total of 5214 particles were used for 2D analysis. After two initial rounds of class averaging with 200 classes, 4719 particles were classified into 32 classes by e2evalparticles.py and e2refine.py from the EMAN2 [33] package. The 24 class averages from classes with at least 75 particles and that showed particles with delineated borders are shown here. The number of particles in each class varied between 75 and 273. Scale bar: 12 nm.

magnification of 50000X. Under these conditions the images typically shows Thon rings visible at least $\sim$ 12Å. A random conical data set of 10 tilt pairs, with tilt angles ranging between 420 and 450 are acquired and the particle pairs of around 2200 are picked as shown in (Figure 4.10). The images were recorded F416 CMOS camera with a physical pixel size is 15.6μm and effective pixel size of 10.9μm at the screen level. The images are not binned and so, the final pixel size of the tilt pairs is 2.1Å. The spherical aberration value (Cs) is 2.0. The technique is to take two exposures at a point on the grid, one should be taken at 0^0 tilt or untilted and other at 450 tilted (same place and expecting a overlap between two) . Pick the particles of the both images at the same place, so that they have one to one correspondences and expected to get another view of the same particle in the tilted view. Next step is to do the class averages from all picked particles from 0^0 tilt images and results in assignment of in plane angle parameters to all particles. Then, due to one to one correspondence, stored angle assignment is also assigned to tilted particles w.r.t the untilted particles. Then back projection algorithms or filtered back projection algorithm is used to get 3D structure model.

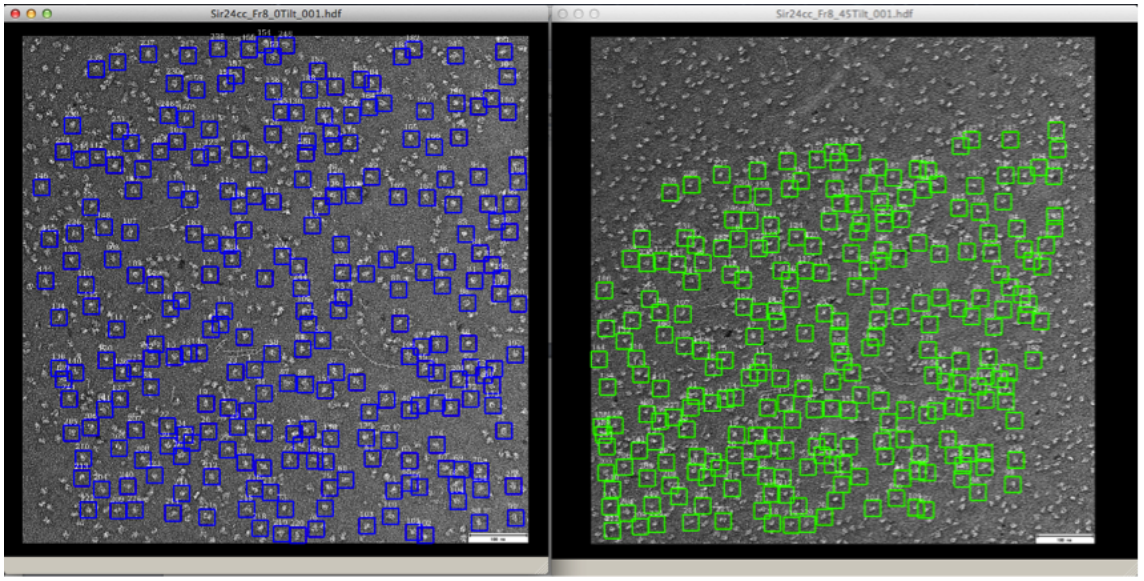


FIGURE 4.10: Gui for Tilt pair picking tool in EMAN2 package

4.7.2 2D and 3D Analysis

In total of 2248 particle pairs are picked and as explained in 2.4 the untilted particles are classified using the 2D classification tool in EMAN2 [33]. A total of 31 classes are generated for heterogeneity analysis as shown in (Figure 4.11A). As explained in the 2.3 the parameters for MSA need to be tuned for every project according to the quality of the data set. For SIR 24 mutant complex as shown in the MALDI-TOF results (Figure 4.4) shows that the sample is a mixture of three different populations. The

decision to choose the number of classes and other parameters makes it easy by keeping into memory of the particle variations during data collection and particle picking. A total of 31 classes is generated with refinement iterations of 5, averaging iterations of 7 by using 6 alignment references, 8 Eigen images and class keep of 0.8. Applying these parameters to Sir 24 cc complex classes are generated with 100 to 150 particles in a good class average. From these Class averages, 22 good ones are picked to make RCT reconstructions (Figure 4.11B). The reconstructions from class averages 0,01,03,06,12 to 14,16 and 21 shows very low resolution features, noisy and undefined structures as expected already because of its smoothness and lack of details in the densities from the class averages in (Figure 4.11A). The reconstruction shows that there is a possibility of three populations in the samples. After carefully viewing all the models in Chimera, the models are classified into three groups, which is also supported by MALDI-TOF results (Figure 4.11C,D,E]. From all the reconstructions, the 3-D models 04, 09 classes are not used since it does not match to any group.

4.7.3 Computing RCT reconstruction form each class average

The RCT reconstruction for each selected class average is calculated and analyzed as shown in the 4.7.2 [33]. To calculate these reconstructions, e2rct.py tool from EMAN2 package is used [33]. This tool needs lot of parameters to calculate 3-D model, which are monitored and tuned for SIR complex dataset. Here are some parameters to be tuned:

1. The first two parameters are Stage tilt and Tilt axis, the former value is the angle between tilt pairs and SIR complex dataset was collected at 450 and the later is calculated during tilt pair picking tool e2rct.py.

2. Bad class averages are rejected and then at least 10 particles are requested to calculate the reconstruction, otherwise a decent 3D volume cannot be generated.

3. Even though the untilted particles are aligned and centered to calculated the 2D class averages, tilted particles are also needed to be centered and aligned by calculated an average and aligning by translation to this average. This process iterated for 6 iterations (user defined). A maximum translational shift of 10 pixels is allowed during the alignment.

4. Averaging all the RCT 3-D models by aligning to each other to produce a optimal one 3-D model Figure 4.12 to average out the missing information termed as “missing cone”. Here due to non-availability of 3-D reference structure the structures are aligned to each other.

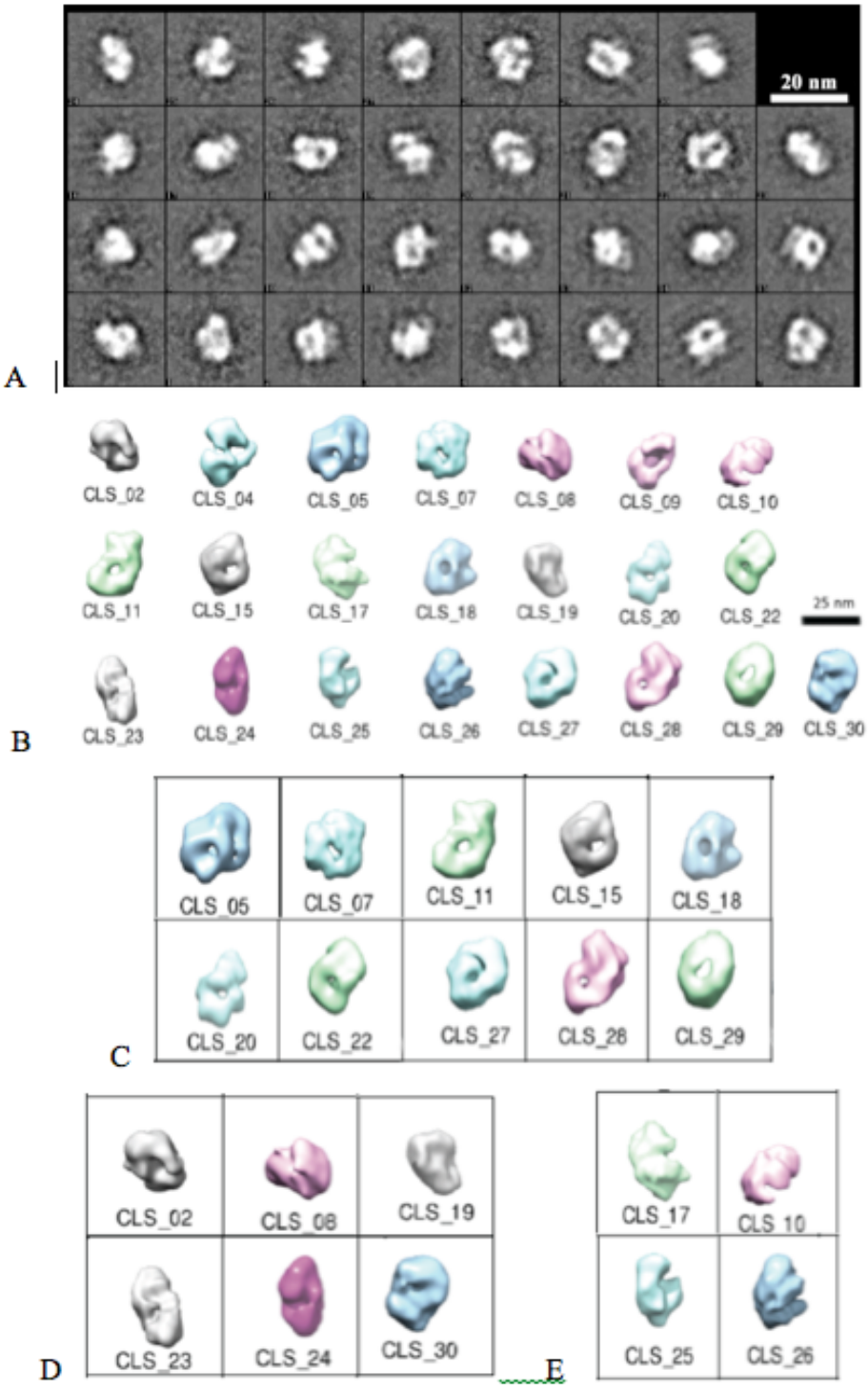


FIGURE 4.11: (A)Class averages of 2248 untilted images. (B)RCT reconstructions from selected 22 best class averages. Manually these reconstructions are classified into three groups according to shape and sizes (C)First Group → classes 05,07,11,15,18,20,22,27,28,29 (D)Second Group → classes 02,08,19,23,24,30 (E)Third Group → classes 17,10,25,26

5. Symmetry is not applied and also due expected low resolution reconstruction the alignment fineness in 3-D of 100 is applied.

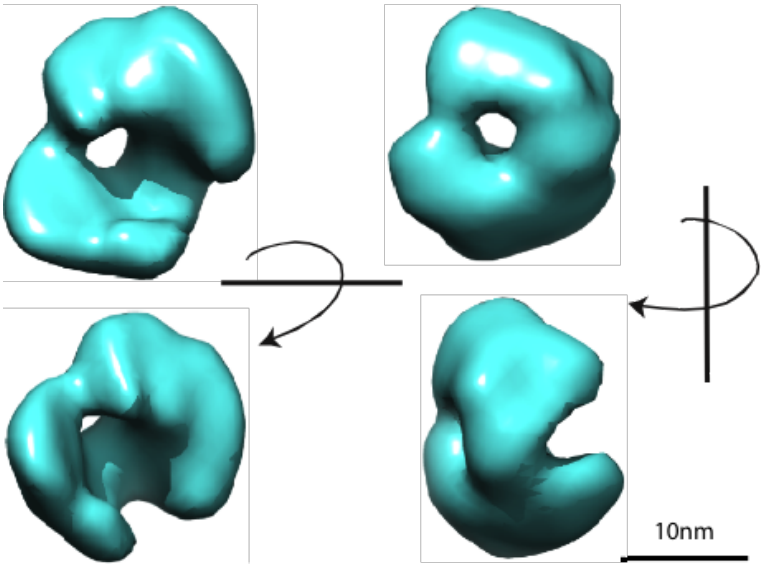


FIGURE 4.12: Final RCT 3-D reconstruction after 3-D alignment and averaging the 11 RCT reconstructions from Group (Figure 4.11c

4.8 Conclusion from EM Analysis of the SIR complex basic units

The Grafix technique seems the best approach for structural study of Sir complex using electron microscopy and single particle analysis. However the class averages showed variability in shapes, indicating heterogeneity in the protein preparations. Systematic structure determination of the SIR Complex by parts is the method to analyse higly heterogenous sample. Once the optimal model determined, the Handedness has to be determined with the help of tilt pairs. The class averages are calculated from the Sir2-Sir4 tetramer Figure 4.13A and the Sir2-Sir4-AB dimer Figure 4.13C datasets and by observing their sizes shows a clear size difference between Figure 4.13B and Figure 4.13D. The Sir2-Sir4-AB dimer averages shows the particles have a diameter of 9.5-15 nm and a roundish shape, whereas the Sir2-Sir4 tetramers seem to be composed of two similarly units with similar shape to Sir2-Sir4-AB as predicted for a tetramer. The angle between the two Sir2-Sir4 dimer units Figure 4.13E seems to vary between the averages. The Sir3-dimer particles are between 7.5 and 9nm wide Figure 4.13F. The semi-automated tool proposed here was proven as a powerful technique to analyze near to 100000 particles. Random Conical Tilt method is a proven method to analyze the heterogeneity analysis and for Sir 24 mutant sample the RCT analysis shows similar results to MALDI-TOF

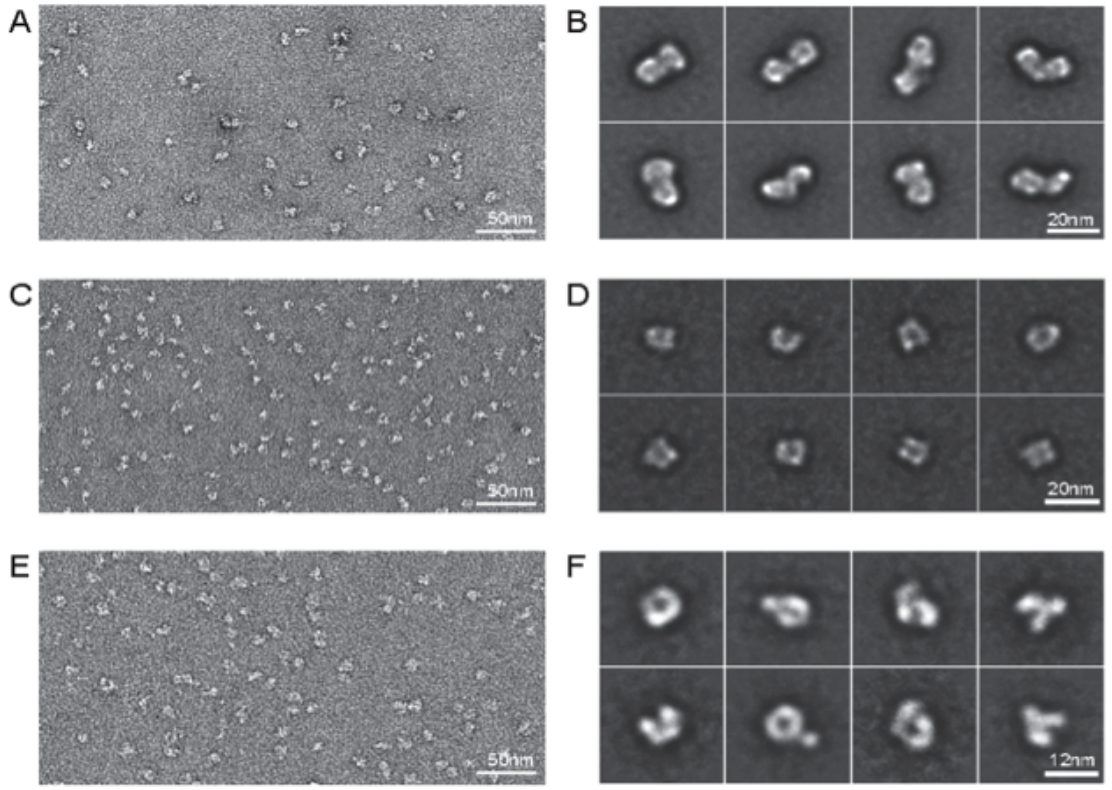


FIGURE 4.13: EM analysis of Sir24 complexes and Sir3 homo dimer

analysis. The poor quality of RCT 3-D models is due to recording the tilt pairs at very low tilt angle difference 45^0, which results in missing of structure factors by 50%. Due to lack of high-resolution structure it is difficult to explain the biological process of Silencing by SIR complex proteins precisely.

Chapter 5

Structure and assembly of the mouse ASC filament by combined NMR spectroscopy and cryo-electron microscopy

Lorenzo Sborgi[1+], Francesco Ravotti[2+], Venkata P. Dandey[3+], Mathias S. Dick[1], Adam Mazur[1], Sina Reckel[1,6], Mohamed Chami[3], Sebastian Scherer[3], Anja Böckmann[4], Edward H. Egelman[5], Henning Stahlberg[3], Petr Broz[1*], Beat H. Meier[2*] and Sebastian Hiller[3,6*]

1. *Biozentrum, University of Basel, Klingelbergstrasse 70, 4056 Basel, Switzerland*

2. *Laboratorium für Physikalische Chemie, ETH Zurich, Vladimir-Prelog-Weg 10, 8093 Zurich, Switzerland*

3. *Center for Cellular Imaging and NanoAnalytics, Biozentrum, University of Basel, Mattenstrasse 26, 4058 Basel, Switzerland*

4. *Institut de Biologie et Chimie des Proteines, 7 passage du Vercors, 69367 Lyon, France* 5. *Department of Biochemistry and Molecular Genetics, University of Virginia, Charlottesville, VA 22908, USA* 6. *Current address: EPF Lausanne, Switzerland*

+equal contribution *Corresponding authors: Petr Broz, petr.broz@unibas.ch; Beat H. Meier, beme@phys.chem.ethz.ch; Sebastian Hiller, sebastian.hiller@unibas.ch;

The following section has been formatted for submission in April 2015. The results shown here are initial version of the manuscript and My contribution to this study was sample preparation for cryo-EM, image recording, data processing, model building and data interpretation.

Abstract

Inflammasomes are multi-protein complexes that control the innate immune response by activating caspase-1 and cytokines of the IL-1 family in response to invading pathogens and endogenous triggers. Assembly of inflammasomes is receptor-specific and induced by activation of a receptor protein, usually a NOD-like receptor (NLR) or PYHIN protein family member. For many of these receptors, the connection to caspase-1 is mediated by the adapter protein ASC (Apoptosis-associated speck–like protein), which consists of two domains. The C-terminal caspase-recruitment domain (CARD) is involved in the recruitment of the caspase, while the self-association of the N-terminal PYRIN domain (PYD) forms a large filament as molecular scaffold of these receptor-specific complexes. Here, we combine three biophysical techniques – solution and solid state NMR spectroscopy, and cryo-EM – to infer structure and filament formation of mouse ASC in vitro at atomic resolution. The complementary information of cryo-EM and solid state NMR is combined in a single structure calculation that renders the atomic resolution structure of the ASC filament. Chemical shifts perturbation of ASC-PYD upon assembly monitor the specific binding interfaces of PYD association, as confirmed by structure-based mutagenesis in vivo. Importantly, NMR experiments show the structural integrity of the PYD assembly within the full-length filament, as well as the mobility of the CARD domain relative to the assembled PYD core. The three-dimensional structure of the mouse ASC-PYD filament is highly similar to the recently determined human ASC-PYD filament, suggesting evolutionary conservation of the ASC-dependent inflammasome.

5.1 Introduction

The innate immune system is characterized by its ability to rapidly detect and reduce different types of pathological agents at minimal concentrations. This is achieved through the recognition of specific pathogen- and danger associated-molecular patterns (PAMPs and DAMPs) by specific, germ-line encoded pattern-recognition receptors (PRRs) that control conserved signaling pathways and induce inflammatory and anti-microbial defenses. A subset of cytosolic PRRs responds to PAMPs by initiating the assembly of cytoplasmic macromolecular complexes known as inflammasomes; a central step for the regulation of the inflammatory cascade leading to the activation of caspase-1, and the proteolytic maturation of pro-interleukin (IL)-1β and pro-IL-18, the final mediators of inflammation and

cell death. The correct structural assembly of inflammasome complexes is critical for the host defense and malfunctions of its molecular constituents are related to major human diseases including cancer and autoimmune syndromes[103].
Inflammasome signaling initiates with the activation of dedicated sensor proteins of the PYHIN family (pyrin domain-(PYD) and hemopoietic expression, interferon-inducibility, nuclear localization domain-(HIN) containing), such as AIM2 and IFI16, or through members of the NLR family (nucleotide-binding and oligomerization domain and leucine-rich-repeat containing), also known as NOD-like receptors [103]. While PYHIN receptors are activated through the interaction of one or two HIN domains with specific ligands such as viral dsDNAs, NLR receptors sense various chemically and structurally different types of molecules, making its activation pathway intricate and still poorly understood. Structurally, the NLRs are characterized by a tripartite architecture defined by a C-terminal leucine-rich domain (LRR), a central nucleotide binding domain NBD (known as NACHT or NOD domain) and an N-terminal effector domain which can be a PYD leading to so-called NLRPs (e.g. NLRP1, NLRP3 and NLRP6), a CARD domain (caspase-recruitment-and-activation-domain) like NLRC4 or a BIR domain (Baculovirus IAP repeat)[104]. Upon activation, the self-association of the receptor is required for the recruitment of procaspase-1 through direct interaction between CARD domains. However, the fact that most inflammasome receptors feature a PYD implies as a preliminary step the recruitment of the bipartite adaptor protein ASC (Apoptosis-associated speck–like protein) that contains both a PYD and CARD. Thus, the protein ASC plays a key role in inflammasome constitution since its C-terminal CARD domain mediates the recruitment of inflammatory caspases while its N-terminal PYD, upon interaction with the receptor PYD, oligomerizes to form the so-called ASC speck, a large filament assembly that acts as a structural platform of the whole inflammasome (Figure 5.1A, B)[105]. Subsequently, procaspase-1 is recruited to the inflammasome, resulting in its autoprocessing and the formation of the catalytically active heterotetramer of cleaved p10 and p20 subunits.
In the last decade, the in vivo observation of the association of ASC into large speck complexes in response to different inflammasome stimuli increasingly focused the attention on the functionality of this protein to unravel inflammation processes. The possibility of studying ASC in vitro using recombinant approaches has highlighted ASC as an attractive target for performing biological and structural studies. Thereby, the structures of the human ASC pyrin domain only (ASC-PYD) and ASC full-length (ASC-FL) by nuclear magnetic resonance spectroscopy (NMR) [106, 107] have opened the possibility for studying the interaction

mode of PYDs and inflammasome assembly in solution [108–110]. Nevertheless, the strong tendency of ASC to oligomerize into large, insoluble molecular aggregates limits the applicability of solution NMR techniques precluding until now the direct observation of individual atoms in this large macromolecular complex. Solid-state NMR spectroscopy under magic angle condition has the potential to overcome this caveat, as demonstrated by structural studies on repetitive filaments, including prion protein fibrils, and injectisomes [111, 112]. The successful applications of solid-state NMR spectroscopy to study the molecular organization of different type of filaments have pointed this technique as an extremely powerful tool for structure characterization of large insoluble fibrous protein aggregates [113–115](11-13).

Recently, the cryo-electron microscopy (cryo-EM) images of the reconstituted human ASC-PYD filament delivered the first structure of the ASC speck [116]. The human ASC-PYD structure provided insights into the mechanisms of homo- and hetero-PYD associations in response to both AIM2 and NLRP3 receptors interaction proposing an unified mechanism of formation for ASC-dependent inflammasomes. Due to the extensive diversification of inflammasome complexes within mammalian species, structural studies of inflammasomes from other mammalian species are of great interest. The close evolutionary relation to humans makes mouse models a method of choice to characterize the immune response and the role of inflammasome components.

In this work, solution and solid state NMR spectroscopy together with cryo-EM are combined in a hybrid structure determination procedure to achieve atomic details of the structure and formation of the mouse ASC filament in vitro. The resulting high-quality structure allows the characterization of ASC at the atomic level within the filament and definition of the interface between the different subunits that are critical for the filament assembly. Additional solid state NMR measurements allow the experimental examination of flexible parts of ASC-FL filament form. Solution NMR spectroscopy is used to characterize structural features of the soluble form of mouse ASC-PYD in isolation. The chemical shift comparison monitors conformational changes in ASC-PYD upon filament formation and identifies the specific residues involved in the binding interfaces. Structure-guided mutagenesis in living cells confirms the network of interactions that are essential for the integrity of the filament and for the propagation of ASC-dependent inflammasome signaling. Overall, our studies provide an atomic level observation of the structure and formation of ASC filament and highlight the immense potential of combined techniques for structural studies of large macromolecular assemblies.

5.2 Results

5.2.1 ASC filament reconstitution *in vitro*

To structurally characterize the assembly of the ASC-dependent inflammasome, the mouse ASC-FL protein and mouse ASC-PYD domain only were expressed and purified from E. coli. Then *in vitro* filament reconstitution was achieved using either of two alternative protocols, initiated from a chaotrope-denatured or an acid-denatured state. In both cases, the ASC-FL and ASC-PYD showed high propensity to assemble into large filaments that were visualized using negatively stained electron microscopy images (Figure 5.1C, D). Oligomerization of mouse ASC-FL and mouse ASC-PYD both led to the formation of long, well-defined single ASC filaments, with typical length in the range of 500–2000 nm. These filaments were organized in larger aggregates of variable size that were suitable for detailed exploration using high-resolution solid-state NMR spectroscopy. The atomic structure of filaments can depend on the assembly conditions in vivo and in vitro and accordingly, the highly polymorphic nature of amyloids has received a lot of attention, for example in the case of the β-amyloid peptide [115]. To characterize the homogeneity of the filaments in our sample preparations, we used solid-state NMR spectroscopy, which is highly sensitive to even small changes in the atomic order of the molecular structural ensemble. The overlay of 2D [13C,13C]-correlation spectra of the ASC-FL and ASC-PYD showed an overlapping set of sharp cross peaks, characteristic of a single predominant structure and thus a nearly identical filament morphology (Figure 5.1E). The perfect overlay of these spectra shows that the isolated PYD domain forms exactly the same filament structure as when it is part of the full-length protein. The ASC-PYD alone thus forms the scaffold of the ASC filament, while the CARD domain is not strictly required for the filament core and does also not perturb the PYD filament conformation. Furthermore, since in these cross-polarization based ssNMR spectra only rigid parts of the filament assembly give rise to strong signals, the absence of additional strong signals in the full-length protein relative tot eh PYD domain only shows that the CARD domain is flexible relative to the filament core. Consistently, while mouse ASC-PYD filaments feature a smooth surface in negative stain EM, filaments formed by mouse ASC-FL have a rough surface, possibly due to the ASC-CARD being exposed at the outside of the ASC filament (Figure 5.1C, D). These data directly indicate that the two domains of mouse ASC, which in human ASC have been shown to tumble independently in solution [106], are independent also in the filament form.

5.2.2 Solid-state NMR spectroscopy of the mouse ASC-PYD filament

The excellent quality of 2D solid-state NMR spectra of the ASC-PYD filament encouraged us to attempt a high-resolution structure determination of the filament. Based on the observation that the ASC-PYD domain is the minimal structural requirement for the formation of native-like speck filaments, we studied the structure of filaments of the mouse ASC-PYD without the CARD. For the assignment of backbone and side chains, a series of ssNMR experiments were recorded on uniformly 13C- and 15N-labeled filaments under magic angle spinning (MAS) conditions. Spectra were well-resolved and allowed sequence-specific resonance assignments of the backbone and amino acid side chains (Figure 5.2B, C, S2). The backbone assignment for the segment residues 3–84 was complete and in this segment, also most of the side chain carbon atoms were assigned. The secondary chemical shifts allowed a direct determination of the location of secondary structure elements. Mouse ASC PYD in its filament form features six α-helices at positions 3–14, 17–29, 41–46, 49–59, 62–76, and 80–84. The last helix presumably is extending until residue 89, but assignments were ambiguous for the segment 85–89.

5.2.3 Cryo-electron microscopy of the ASC-PYD filament

Homogenous and well-ordered ASC-PYD filaments for cryo-EM studies were obtained by polymerizing ASC-PYD by dilution from denaturing solution with aqueous buffer. Optimization of the refolding protocol by observation with cryo-EM shows that these filaments are well-ordered (Fig. S2). Image collection was done manually on a Titan Krios using a CMOS based back-thinned K2 direct electron detector. The averaged power spectrum of multiple individual segments from these cryo-EM images shows a clear meridonial spot at $1/14.2\text{Å}^{-1}$, corresponding directly to the axial rise in Fourier space. Image processing with the Iterative Helical Real Space Reconstruction (IHRSR) method [25] yielded a final electron density map with approximately 4Å resolution. The ASC-PYD filament is a hollow fibre with inner and outer diameters of 20Å and 90Å, respectively. The polar filament has a C3 point group symmetry with 53^0 right-handed rotation and 14.2Å axial rise per subunit, after correcting for a mean out-of-plane tilt of approximately 12^0.

5.2.4 Combined structure calculation of the mouse ASC-PYD filament

The cryo-EM density map of ASC-PYD filament was combined with solid state NMR data towards a joined structure calculation (Fig. 2A). An initial model of the ASC-PYD filament was build into the cryo-EM density by placing an unfolded ASC-PYD polypeptide chain manually into the cryo-EM density. Helical regions were defined de novo, based on the backbone dihedral angles determined by TALOS+([117]) from the solid state NMR chemical shift values. The resulting model was refined under symmetry enforcement, using as constraints the dihedral angle values and hydrogen bonds from solid state NMR data and cryo-EM data using XPLOR-NIH [118]. The resulting structural ensemble was then used together with the sequence-specific resonance assignment as input for the automatic peak assignment of 2D [^{13}C,^{13}C]-CHHC and [^{13}C,^{13}C]-PARR spectra using the CANDID algorithm[115]. Importantly, in this cycle of refinement, CANDID can interpret 91% of all observed cross peaks, giving rise to a total of 674 distance constraints (Figure 5.9, Table 5.1). A final XPLOR-NIH structure calculation was repeated with cryo-EM derived structure factors, TALOS+-derived dihedral angle and ssNMR distance restraints under symmetry enforcement. The resulting three-dimensional structure calculation of the ASC-PYD filament integrates the overall symmetry and spatial organization of the subunits inside the filament (as the degree of angle rotation and axial rise per subunit) together with 674 internuclear distance constraints, and dihedral backbone angle constraints for 70 residues (Table S1). The ensemble of 10 lowest energy conformers features a backbone rmsd of 0.14Å and a heavy atom rmsd of 0.48Å (Figure 5.2E, Table 5.1). The ensemble does not show significant deviations from covalent geometry and is well defined by the cryo-EM and NMR data (Figure 5.2D, F, G).

5.2.5 Structure of the ASC filament

The mouse ASC-PYD filament forms a triple-stranded right-handed helical oligomer, where each PYD subunit interacts with the adjacent six PYD modules through three different asymmetric interfaces (Fig. 3A, B). ASC filament strands propagate through the interface type I defined by residues belonging to helices 1 and 4 (type IA) and helix 3 (type IB) of the next subunit. The contact of the central strand with the lower helical layer defines the interface type II interface where the loop between helices 4 and 5 and the central part of loop before helix 3 (type IIA) interact with residues at the corner of helices 5 and 6 of the next subunit.

Finally, the residues from the end of helix 1 and the following short loop (type IIIA) interact with helix 3 (type IIIB) defining the interface type III that mediates the contact of the central strand with the upper helical layer.

5.2.6 Assembly of ASC-PYD into the filament form

To characterize the transition from monomeric PYD to PYD filaments, we characterized the structural features of the soluble form of mouse ASC-PYD by solution state NMR spectroscopy. Mouse ASC-PYD was not soluble at physiological pH solution, in agreement with previous reports (5). However, low pH condition improved the solubility, leading to properly folded monomeric species as indicated by the signal dispersion and narrow line widths of the 2D [^{15}N,^{1}H]-HSQC spectrum (Figure 5.10). Sequence-specific resonance assignments of $^{1}H^{N}$, 15^{N}, $^{13}C\alpha$ and $^{13}C\beta$ nuclei confirmed the position of the six α-helices at positions 3–14, 17–29, 41–46, 49–59, 62–76, and 80–84 as the secondary structure elements of mouse ASC-PYD in solution. The mouse ASC-PYD thus shares the general features of the human homologue and other known PYD structures characterized by six helices and the characteristic long loop between helices 2 and 3 (Figure 5.4A, 5.11). The available chemical shift assignments in both the soluble and filamentous form allow a direct chemical shift comparison to monitor conformational changes upon filament formation and thus to define the subunit contact sites. Isotropic ^{15}N, $^{13}C\alpha$ $^{13}C\beta$ chemical shift and a comparison of the TALOS+ (18) derived φ and ψ torsion angles confirms that all secondary structural features are retained in the ASC-PYD during the filament assembly and thus the absence of major structural rearrangements (Fig. S3). Importantly, the chemical shift comparison between soluble and filamentous form directly identifies residues defining crucial subunit interfaces. These residues show significantly large chemical shift variations, defining specific surface areas on ASC-PYD, and thus independently confirm the three asymmetric interfaces characteristic of the filament architecture (Figure 5.4C). Thereby, the interface type I is formed by interactions between residues of opposite charge. Solvent exposed side chains of residues Asp6, Asp10 and Glu13 from helix 1 together with residues Asp48 and Asp54 from helix 4 create a defined negatively charged surface faced to the positively charged residues Arg5, Lys22 and Lys24 from helix 2 of the following subunit. Structural details highlight others solvent exposed side chains like Lys 21 and Lys 26 from helix 2 that further contribute to create a positively charged counterpart. Moreover, the hydrophobic

network of interactions defined by Leu9, Met25, Leu28 and Val30 play an important role in this surface (Figure 5.4D). Hydrophobic interactions seem to be the major driving force defining the Interface type II. The Tyr59 and Tyr60 from the loop before helix 5 and Leu78 from the loop before helix 6 characterize the interaction between each filament strand and the lower helical layers. However, the observed contribution of further charged and uncharged side chains suggests a more heterogeneous network of interaction. Finally, the charged residues are characterizing the interaction type III required for the contact of each filament strands with the upper helical layers. Negatively charged side chains of Glu13 and Glu19 from end of helix1 and beginning of helix 2 respectively interact with Arg38 and Arg41 from the long loop before the helix3. Overall, the interaction modes of ASC-PYDs define the assembling process of the filament architecture. The filament formation is structurally characterized by primary filament strands propagation that happens through the large complementary charge surfaces of the interface type I. Individual helical strands formation is at the same time supported by the upper and lower layers interaction through additional interfaces II and III stabilizing the final star-shaped helical architecture.

5.2.7 Structural conservation of the ASC-PYD filament

A comparison of the mouse ASC-PYD filament structures established in the present work with the human ASC-PYD filament (PDB 3J63 [116]) shows that the spatial assemblies are highly similar (Figure 5.5A). The human and mouse ASC-PYD subunits as part of the filament assembly overlay with a backbone RMSD of 1.1Å and also the three-dimensional arrangement of the subunits towards the filament structure shows the same overall arrangement. This agreement strongly suggests a functional conservation of the ASC filamentation mechanism as part of the innate immune response system in mice, men and possible other mammalian species. This conservation is furthermore highly supported by the fact that the network of key residues required for the filament propagation is well conserved among species (Figure 5.5B). Notably, the high level of residue conservation (57% identity over 90 residues) among different mammalian ASC-PYDs suggests a shared mode of self-association based on residue interactions dominated by electrostatic contacts and by specific hydrophobic side chains as described for the mouse ASC-PYD. Since the ASC-PYD is activated by interactions with upstream PYDs from receptors, it is highly interesting to investigate if the emerged cluster

of residues is additionally required for these interactions. Amino acid conservation analysis has thus been further evaluated for mouse ASC-PYD, AIM2-PYD, NALP3-PYD, and NAL10-PYD (Figure 5.5C). Interestingly, the key residues for the ASC-PYD filament propagation are only weakly conserved in the receptor PYD domains (6% identity over 90 residues). Among the mouse ASC-PYD sequence, only Leu12 from helix 1 and residues Leu20, Lys21, Phe23 and Lys24 from helix 4 are shared in the PYD receptors, in strong contrast to the high conservation of residues among mammalian ASC-PYDs. This finding thus suggests that seeding of the ASC-PYD filament formation by the upstream PYDs of the NOD-like receptors requires additional residues with respect to those involved in the homotypic PYD–PYD interactions.

5.2.8 Dynamic and flexibility of the CARD domain

Solid-state NMR spectra under magic angle spinning conditions can generally be employed with different schemes to transfer polarizations between nuclei. Experiments based on cross-polarization (CP) techniques filter for rigid parts of the assemblies, whereas INEPT-based experiments monitor flexible parts of the molecular assemblies. Our initial CP-based experiments of the ASC-FL filaments have established that the rigid parts of the filament are entirely formed by the PYD domain and that the CARD domain is a flexible part of the filament arrangment (Figure 5.1). Consequently, INEPT-based experiments were used to obtain spectral information on the conformation and dynamics of this domain. INEPT-based 2D correlation experiments of full-length mouse ASC filaments show a set of backbone amide correlation cross peaks, with chemical shift values in the random coil region (Figure 5.6A). These peaks must arise from the CARD domain, since corresponding spectra of mouse ASC-PYD are devoid of peaks (data not shown). Furthermore, the narrow dispersion of amide proton chemical shifts indicate that the CARD domain populates a conformational ensemble of flexibly unfolded structures in fast equilibrium, similar to a random coil ensemble. Consistently, 2D INEPT-based [^{1}H,^{13}C]-correlation spectra of full length mouse ASC filament feature also no significant chemical shift dispersion, resulting in the observation of few resonance correlations at random-coil chemical shift position (Figure 5.6B). These peaks result from the superimposition of resonances of the same kind of amino acid side chains and confirm the conformational averaging of side chain rotamers in the polypeptide chain. In our preparations of the mouse ASC

filament, the CARD domain is thus flexibly unfolded and highly mobile relative to the folded filament core (Figure 5.6C).

Effect of single points mutations on ASC signaling Structure-based mutagenesis studies in living cells are important tools to confirm the interaction network and to evaluate the integrity of the ASC-PYD filament and relative effects on downstream signaling events, i.e. ASC speck formation, caspase activation and the induction of cell death and cytokine secretion. Previous studies using site-directed mutagenesis of ASC residues, opted for overexpression of fluorescently-tagged ASC-FL or ASC-PYD, which results in spontaneous formation of ASC specks or filaments, even in absence of receptor mediated activation. However, since different receptors and stimuli can result in the formation of different subtypes of inflammasome complexes, such assays fail to confirm the physiological relevance of ASC mutations for inflammasome activity. To test the physiological effects of ASC mutations on inflammasome signaling in response to activation of different NLR and PYHIN receptors and to avoid over-expression artifacts, we reconstituted immortalized mouse Asc-/- macrophages with endogenous levels of WT and mutant versions of mCherry-tagged full-length ASC. Filament assembly occurs by single strand propagation and simultaneous interactions with proximal layers. Thus, it appears most likely that specific charged residues that belong to interface type I are of primary importance for filament reconstruction and thus suitable for testing mutation effects. Based on the mouse ASC-PYD sequence alignment (Figure 5.5A) and previously reported mutagenesis studies on the human homologue (19), the effect of K21A and K26A mutations on the filament morphology and inflammatory signaling was studied (Figure 5.7). As expected, the induction of cell death and IL-1β secretion upon activation of the PYD-containing inflammasome sensors NLRP3 and AIM2 by ATP treatment or DNA transfection, respectively, was abrogated by mutations in K21 and K26 when compared to levels induced by WT ASC (Figure 5.7A,D). Consistent with the deficiency in signaling, no ASC specks could be detected in macrophages expressing these two single amino acid mutants (Figure 5.1B,C,E). A distinct feature of the NLRC class of inflammasome receptors is that they feature a CARD domain instead of a PYD, and thus could in theory directly recruit pro-caspase-1. Indeed, the induction of cell death upon stimulation of these receptors, such as NLRC4 and murine Nlrp1b, is known to be ASC-independent (20, 21). Nevertheless, ASC was shown to be required for pro-caspase-1 processing and efficient IL-1β secretion upon activation of NLRC4 and Nlrp1b (21, 22). Consistently, we observed that K21A and K26A mutations had no effect on the induction of cell death in macrophages infected with Salmonella

enterica serovar Typhimurium, a potent activator of NLRC4 (Figure 5.7F). However, the mutations significantly reduced IL-1β release and completely abolished ASC speck formation during Salmonella infection (Figure 5.7 F,G), in line with previous reports that show the importance of functional ASC on these aspects of NLRC4 biology[119]. Thus, the disruption (or reduction) of filament formation upon K21A and K26A mutations confirms the importance of the precise balance of charged residues of the interface type I and it suggests that the mechanism of filament assembly proposed for the in vitro reconstruction is extendable to the in vivo conditions. Moreover, our results show that mutations in the type I interface prevent ASC speck formation and inflammasome signaling during activation of the three best studied inflammasome receptors NLRP3, AIM2 and NLRC4. This indicates that, independently of which ligand and receptor induce inflammasome activation, the architecture of the ASC filaments remains conserved and consequently the ASC speck remains conserved.

5.3 Discussion

The current work reports the determination of the three-dimensional structure of the mouse ASC filament by a combination of atomic resolution data from NMR spectroscopy and cryo-electron microscopy. The ASC filament is a helical arrangement of individual ASC-PYD domains, mediated by electrostatic intermolecular contacts. The ASC-CARD domain is flexibly attached on the outside of the filament core and was found to be unfolded in our experimental conditions. It will be of interest to investigate whether this shifts towards an unfolded conformation is a general property of the CARD domain and what functional role it may have as part of the immune response. For example, this could be a relevant part of the interaction mode with downstream caspases. The combination of solid state NMR restraints with a cryo-EM desnity employed here is one of very few recent examples of this emerging approach. The other published example comprises the type III secretion needle [112, 120]. Importantly, this techniques provide complementary information on both the well-defined and the dynamically unfolded parts of a structural assembly, leading to an overall comprehensive description that extents the power of the cryo-EM or ssNMR alone, and leads to structural descriptions with increased precisions. Individual internuclear contacts and dihedral backbone angles from ssNMR data allow unambiguous determination of the backbone structure de novo and cross-validate individual side chain conformations. These features are of particular relevance in regions where the global or

local resolution of cryo-EM may not allow resolving all structural details of interest, or where the solution structure is not available or substantially different from the structure in the rigid assembly. The comparison of the mouse ASC filament structure with the previously established human ASC structure shows that the molecular architecture of the ASC filament and thus the molecular mechanism of ASC-dependent innate immune response is strongly conserved between men and mice. The filament architecture is thus encoded in the amino acid sequence of ASC and determines the correct quaternary assembly upon inflammatory stimulation. The mouse ASC filament provides an ideal basis for structure-based mutagenesis, as demonstrated with our single point mutation experiments in vivo. Overall, this work provides a platform for combined *in vivo* and *in vitro* studies towards the structural biology of ASC-dependent inflammasomes and their upstream and downstream regulation.

Acknowledgements

We thank Vesna Oliveri, Janine Zankl, and Timm Maier for experimental help and discussions.

5.4 Experimental Procedures

5.4.1 Cloning, expression and purification of ASC-FL and ASC-PYD

The DNAs coding for the full length mouse ASC protein (residues 1–193) and for the PYD domain (residues 1–91) were cloned with a C-terminal six-histidine tag into a PET28a vector under the control of a T7 promoter. Previous study shown enzymes inefficiency for the cleavage of the His-tag at the lower pH condition required for ASC solubility [121]. Thus, an extra GSGSLE linker was introduced at the C-terminus to minimize His-tag effect on protein structures. Both protein constructs were transformed in BL21(DE3) E. coli strains and the proteins were expressed by growing the cultures at 37^0C to an OD600 of 0.8 and by inducing with 1 mM IPTG for 4 hours. ASC-FL and ASC-PYD [U-^{15}N]- and [U-^{15}N,^{13}C]-labelled, were produced using 13C-glucose and 15NH4Cl as the sole carbon and nitrogen sources. The cells were harvested by centrifugation and the pellet was resuspended in 50 mM Phospahate buffer pH=7.5, 300 mM NaCl, 0.1 mM protease inhibitor. The resuspended cells were incubated for 1 h at room temperature with DNase I and then sonicated on ice and centrifugated at 13000 rpm at 40C for 30 minutes. Both ASC-FL and ASC-PYD were not present in soluble fractions thus denaturing conditions were used to solubilize them from the inclusion bodies. The pellet was resuspended in 50 mM phospahate buffer pH=7.5, 300 mM NaCl, 6M guanidinium hydrochloride and centrifugated at 13000 rpm at 40C for 30 minutes. The supernatant was incubated for 2 hours at room temperature with pre-equilibrated Ni-NTA affinity resin (Thermo Scientific) and than passed through a plastic body column for gravity flow purification. The column was washed with 20 column volumes of resuspension buffer containing 20 mM imidazole and the fusion protein was eluted with 3 column volumes of the same buffer with 500 mM imidazole. To avoid aggregation, all the purification steps were carried out at 40C and further addition of 2 mM DTT to all the buffers were used for the ASC-FL purification. An Additional step of purification was used to obtain the monomeric soluble form of ASC-PYD. ASC precipitates at physiological pH condition thus the pH of the elution fraction from Ni column was decreased to 3.8 and dialyzed against 50 mM Glycine buffer pH=3.8, 150 mM NaCl. The protein was further purified on a pre-equilibrated Superdex 75 gel filtration column (GE Healthcare) and it was either used immediately or frozen in small aliquots in liquid N_2.

5.4.2 ASC-FL and ASC-PYD filament formation in vitro

Two protocols for ASC-FL and ASC-PYD filament formation based on pH- or dialysis-step were used leading to filament structures as determined by negative stain EM. For the pH-step based method, the elution fraction from Ni column was initially concentrated to half of the volume ($\backsim$5 ml) using Vivaspin and then diluted in a volume ratio of 1 to 9 with 150 mM acetic acid pH=2.5. The neutral pH condition was achieved by the addition of 3M Tris buffer pH=8 in a volume ratio of 1 to 5. The solution was incubated overnight at room temperature under continuous stirring condition to facilitate filament formation. The solution was centrifugated at 13000 rpm at 40C for 30 minutes yealding a gel-like pellet of ASC filaments that was resuspended in water for a wash step and then transferred into the rotor by centrifugation or stored at 40C. For dialysis based filament formation method the elution fraction of NiNTA purification was dialyzed overnight against 25 mM Phosphate buffer pH=7.5, 100 mM NaCl. The gel-like containing ASC filament were centrifugated, washed and stored as previously described.

5.4.3 Solution NMR spectroscopy

NMR samples were prepared at 0.3 mM [U-^{15}N,^{13}C]-labelled ASC-PYD in 20 mM glycine buffer pH=3.7, 150 mM NaCl, 0.1 mM NaN^3, , 5% D_2O/H_2O. Under these conditions the ASC-PYD remains soluble and monomeric according to NMR line width values. NMR experiments were acquired at 298 K in a Bruker Ascend II 700 MHz spectrometer equipped with a cryogenic triple resonance triaxial-gradient probe. Sequence backbone chemical shift assignments were obtained from the experiments [^{1}5N-^{1}H]-HSQC, 3D HNCACB and 3D C(CO)NH-TOCSY. All the NMR data were processed using PROSA [122] and analyzed with CARA [123].

5.4.4 Solid-state NMR spectroscopy

Protein filaments were packed into rotors using ultracentrifugation. For the assignment, multiple ssNMR spectra were recorded and analyzed.

5.4.5 Cryo-EM microscopy and image reconstruction

ASC-PYD filaments polymerized with buffer solution were applied to glow discharged thin carbon film-coated copper EM grids. The grids were blotted for 1 s

and vitrified by plunging them into liquid nitrogen-cooled liquid ethane with help of FEI Vitrobot MK4. These grids containing ASC-PYD filaments were imaged using a FEI Titan Krios electron microscope operated at acceleration voltage 300 KeV, and recorded using a 4k x 4k K2 Summit direct electron detector (DED) with a back thinned CMOS chip. The images of size 4k x 4k were recorded by operating the DED in electron counting-mode according to the procedure in [124]. In total, 150 images are recorded by operating the microscope in low-dose mode at a nominal magnification of 22,500x with a cumulated dose below $\backsim$20 $e^-/Å^2$ distributed over 30 frames recorded over 5s. The effective pixel size of the images was 1.34Å on the sample level. The software [125] was used to do drift alignment of individual frames automatically by employing drift correction algorithm by, which works in real time while acquiring the images in the microscope[126].

The CTF was determined using CTFFIND3 [127], and the CTF was corrected by multiplying the images with the theoretical CTF (a Weiner Filter in the limit of a low signal-to-noise ratio). Long filament sections were cut from the images using the e2helixboxer routine within EMAN2 [33]. The SPIDER software package was used for most of the image processing [58]. A total of 21,138 overlapping 320 px long boxes, with 1.34 Å/px, were then cut from the long filament boxes, using a shift of 16 px (95% overlap). These segments were initially decimated by a factor of three (to 4.0Å/px) and centered. Subsequent processing involved the IHRSR algorithm [27]implemented within SPIDER, with the imposition of a C3 rotational symmetry about the helical axis. The final reconstruction was corrected for the CTF (which had been imposed twice, once by the microscope, once by us) by dividing by the sum of the CTFs squared. A reconstruction was tried using a 384 px box length but suffered from worse resolution, reflecting the limited long-range order in these filaments.

5.4.6 Structure calculation procedure

An initial model of ASC-PYD filament was built into the cryo-EM density representing 15 monomers. Dihedral backbone constraints for 70 residues were obtained from solid-state chemical shift values using the software TALOS+[117]. In the first step of calculation, a single ASC-PYD subunit was constructed de novo using the program coot [128]. Thereby, the unfolded ASC-PYD polypeptide chain was placed into the cryo-EM density and the helical regions were defined according to the TALOS+ constraints. Well-defined side-chain electron density for the residues F23, R38, Y59, Y60 and Y64 facilitated the correct positioning of the

polypeptide chain into the density map. The structure of the single monomer was then translated into all symmetrically identical positions of the electron density map and the resulting model was refined locally to match the geometrical restraints and the electron density map. The model was further refined against solid state NMR data (dihedral angle values and hydrogen bonds) and cryo-EM data using XPLOR-NIH [115]. For this purpose, the cryo-EM map was placed into an artificial crystal lattice environment and transformed to reciprocal space representation to calculate structure factors. The amplitudes and phases of the structure factors were used as experimental diffraction data for model refinement by crystallographic conjugate gradient minimization, B-factor optimization and simulated annealing refinement. This conversion from real space cryo-EM map to reciprocal space structure factors is necessary for the simultaneous refinement with NMR data. The structure calculation protocol was repeated 100 times and 20 lowest energy structures were selected. The resulting structural ensemble was used as input for the automatic assignment of the solid state CHHC and PARR spectra. The spectra were automatically peak picked and the peak lists were interpreted using the CANDID algorithm in the structure calculation program CYANA [129]. Chemical shift tolerances for CANDID were set at 0.20 ppm for ^{13}C and only the last cycle (cycle 7) of CYANA automatic protocol was applied. The peak lists were converted to upper limit distance restraints assuming a uniform distance limit of 7Å for every peak. The final XPLOR-NIH structure calculation was performed with the cryo-EM derived structure factors, dihedral angle constraints from TALOS+ and the ss-NMR distance restraints from CANDID. 10 lowest energy structures were selected out of 100 structures.

5.4.7 Cell culture

WT and *Asc-/-* immortalized murine bone marrow derived macrophages [119] were cultured in DMEM (Sigma) supplemented with 10% 3T3-MCSF supernatant and 10% FCS (BioConcept).

5.4.8 Retroviral Constructs

Genes encoding mCherry and murine Asc were cloned into a replication-defective mouse stem cell retroviral construct (pMSCV2.2). Site-directed mutagenesis was performed with QuickChange (Stratagene).

5.4.9 Cell lines

Retroviral transduction of iBMDMs with ASC-mCherry and the K21A and K26A mutations was done as previously described [119] and monoclonal cell lines were generated by limited dilution [117]. ASC protein levels were estimated by Western blot using anti-ASC (Genentech) and anti-β-actin (Sigma) antibodies and clones expressing comparable levels of ASC as WT cells were selected.

5.4.10 Inflammasome activation

Immortalized BMDMs were pre-stimulated for 4 h with 100 ng/ml LPS O55:B5 (Invivogen). NLRP3 was triggered by 5 mM ATP (Sigma) for 90 min, AIM2 by transfection of 1 μg/ml poly(dA:dT) (Invivogen) in OptiMEM (Gibco) using Lipofectamine2000 (Invitrogen) for 2.5 h and NLRC4 by infection with log-phase Salmonella enterica serovar Typhimurium SL1344 wild type at an MOI of 10 for 60 min. Cell death was measured by LDH release (Clontech LDH Cytotoxicity Detection Kit) and IL-1b by ELISA (eBioscience)

5.4.11 Confocal microscopy

To assess the formation of ASC-mCherry specks the cells were seeded on glass coverslips and inflammasome activation triggered as described above. To prevent cell death and subsequent detachment from the coverslips, the cells were treated with 25μM Z-Val-Ala-DL-Asp-fluoromethylketone (Bachem) to inhibit caspase activation at the time of inflammasome activation. iBMDMs were fixed with 4% PFA, mounted with Vectashield containing DAPI and imaged using a Leica Point Scanning Confocal "SP8" with either a HC PL APO CS 40x objective for speck quantification or a HC PL APO CS2 63x objective for overview images. Images were acquired using the Leica LAS AF software (version 3) and analyzed using Fiji. For speck quantifications, Z-stacks of five arbitrarily chosen regions per condition and replicate were acquired and the number of cells with specks determined manually.

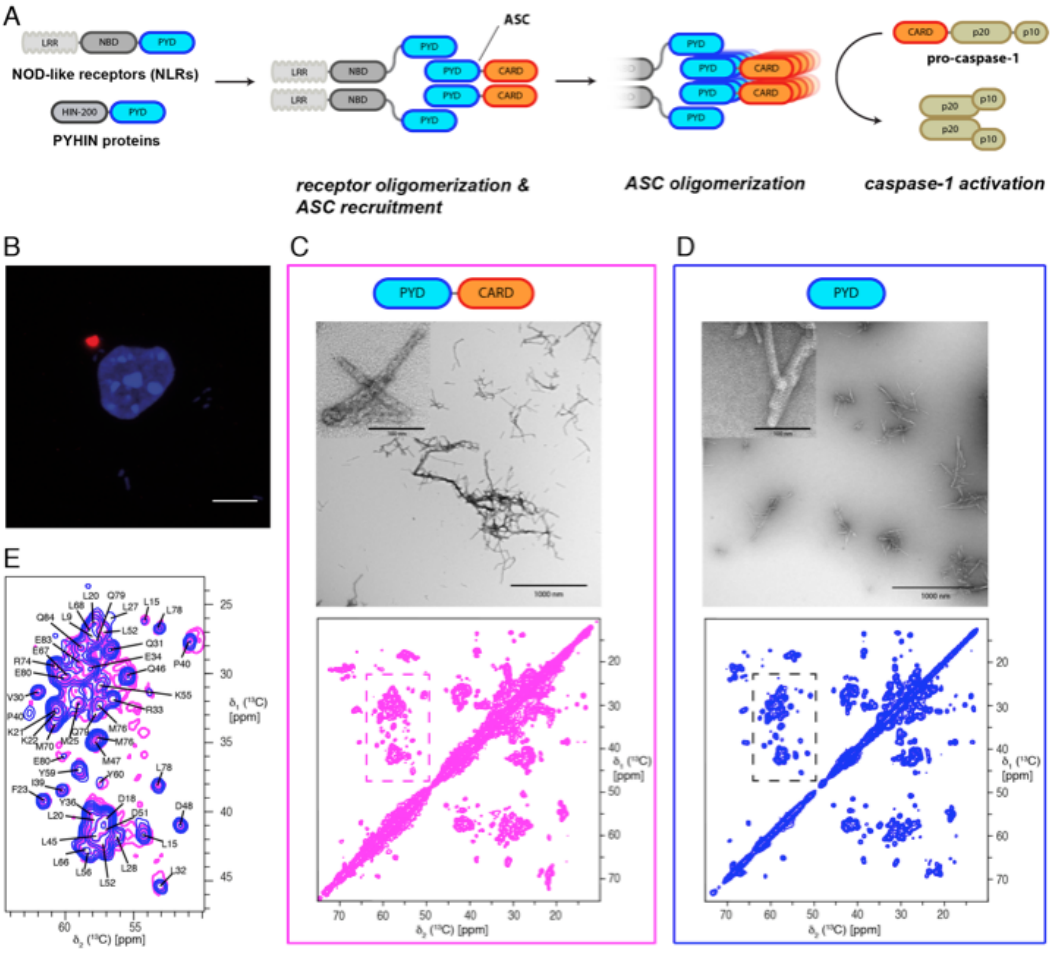


FIGURE 5.1: Structural determinants of the ASC filament formation. (A) Schematic representation of ASC-mediated inflammasome assembly and caspase-1 activation. Detection of specific molecular patterns by NLR or PYHIN family members result in their activation and oligomerization. Activated receptors recruit the inflammasome adaptor ASC, which in turn oligomerizes to ASC filament. The inflammasome effector protease, caspase-1 is activated by this complex. (B) Endogenous ASC specks in primary murine macrophages stained with antibodies for ASC appear as large, macromolecular complexes. Scale bar = 5 μm. (C, D) Morphologic characterization of filament from ASC-FL and ASC-PYD reconstituted in vitro. Negatively stained images of ASC filaments (scale bar 1000 nm) and 2D [13C,13C]-DARR spectra of [U-13C]-labelled filaments. The ssNMR spectra were recorded on a 850MHz spectrometer at 17 kHz MAS conditions. (E) Superimposition of selected region of the 2D [13C,13C]-DARR spectra of the ASC-FL

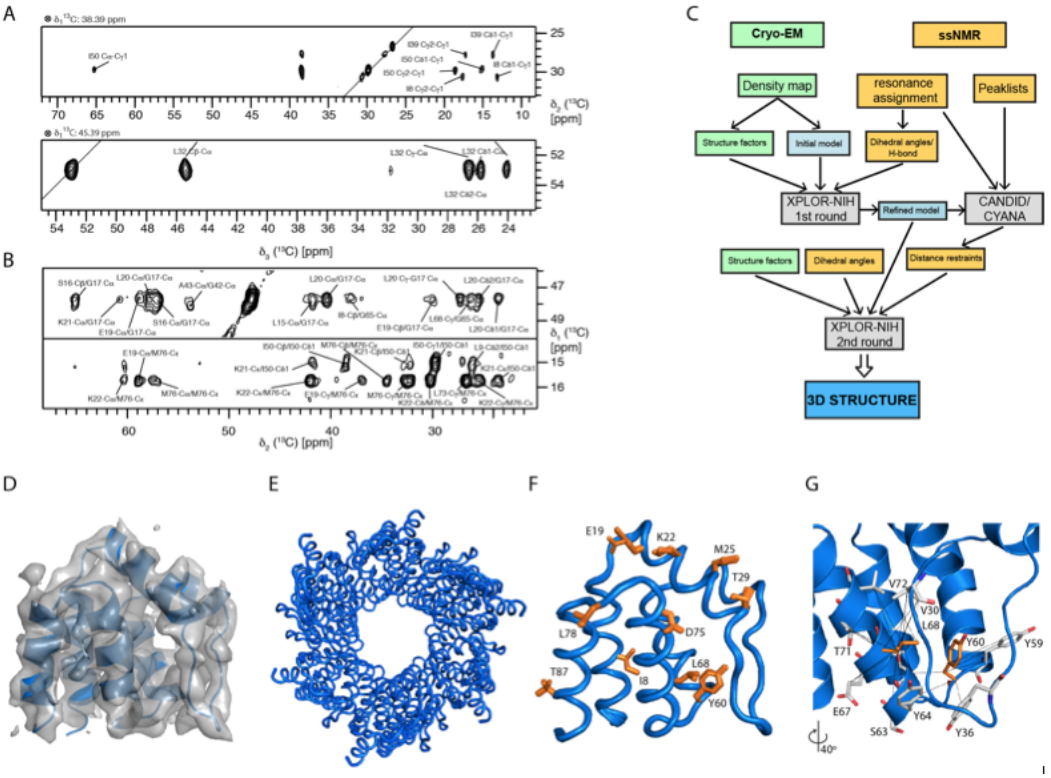


FIGURE 5.2: Structure determination of the ASC-PYD filament by a combination of ssNMR and cryo-EM. (A) Strips from a 3D carbon correlation spectrum used for side chain ^{13}C chemical shift assignments on a uniformly ^{13}C-labelled ASC-PYD filament. The strips were taken at the ^{13}Cα positions of residues Ile50 and Leu32, respectively. (B) Strips from a 2D [^{13}C,^{13}C]-DARR (top) and a CHHC spectrum (bottom). Pairwise contacts as identified with the software CANDID are shown using the one-letter amino acid code and residue number. (C) Flowchart of the *ab initio* structure determination of the ASC-PYD filament including the validation of atomic distance constraints from ssNMR experiments. The contributions from ssNMR and cryo-EM restraints to the hybrid structure calculation procedure are reported with orange and green rectangles, respectively. The preliminary models based on reduced sets of ssNMR and cryo-EM data are shown in light blue. The validated distance peaklists by CANDID are combined with solid-state NMR and cryo-EM data in a single *ab initio* calculation procedure to yield the high resolution structure of the ASC-PYD filament (blue rectangle). See Material and Methods for details. (D) cryo-EM density reconstruction superimposed with the single subunit ASC-PYD structure. (E) Backbone superimposition of the 20 lowest energy conformers of the ASC-PYD filament and (E) single unit within the assembled filament. The conformational ensemble of 10 selected side chains is shown as stick models in orange. (F) Atomic level structural features of a single ASC-PYD unit as part of the filament. The experimentally observed nuclear contacts between Tyr60, Leu68 (orange) and the neighbouring residues (gray) from ssNMR experiments are shown with solid black lines.

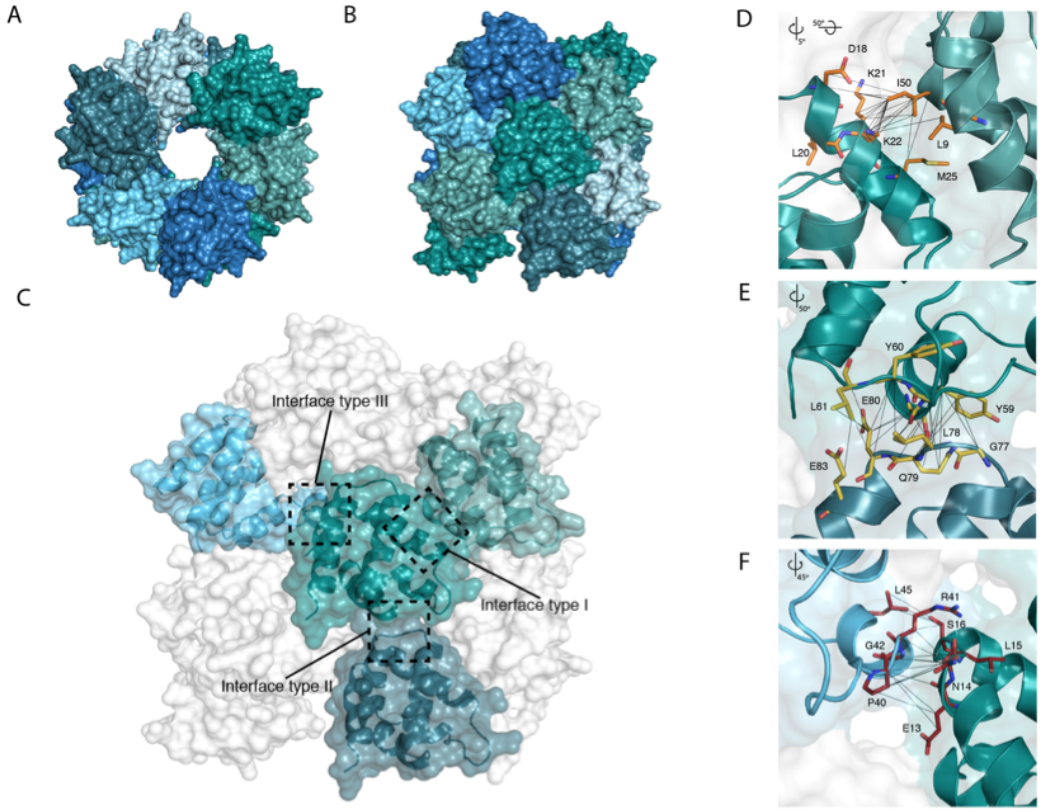


FIGURE 5.3: Three-dimensional assembly of the ASC-PYD filament. (A–B) Top and side views of the ASC-PYD filament in surface representation. The three helical oligomer strands are coloured in blue, teal and dark blue, and single ASC-PYD units within each strand are reported with alternating darker and lighter colour shades. (C) The structures of four ASC-PYD units are reported over the surface representation of the filament using the colour codes previously defined. The three different interaction interfaces involving ASC-PYD units are indicated with dashed square rectangles. (D–F) Detailed view of interaction interfaces I, II and III, respectively. The network of experimentally observed atom pair contacts between residues of different ASC-PYD units is reported with solid black line. The residues defining the interfaces I, II and III are labelled and coloured in dark orange, dark yellow and dark red, respectively.

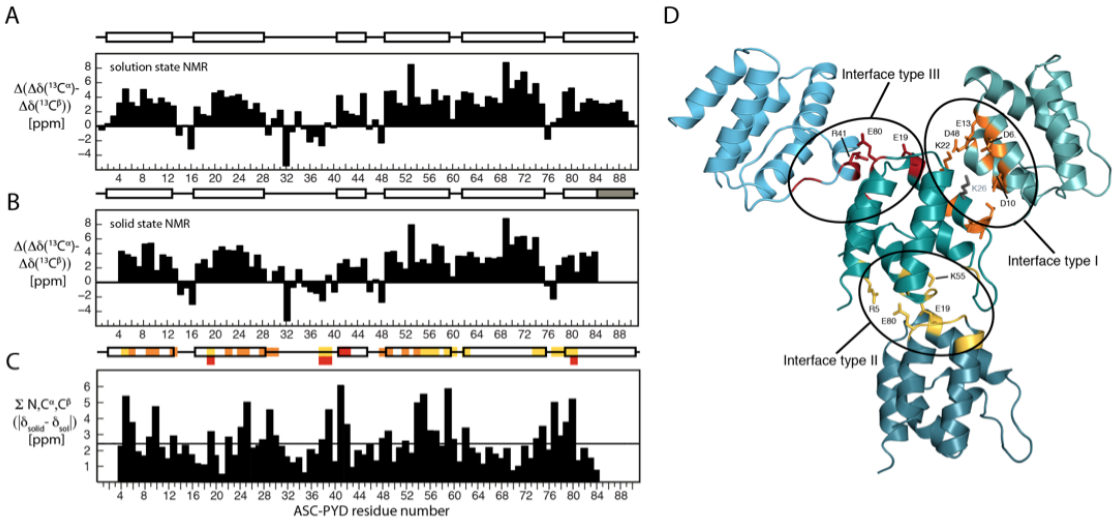


FIGURE 5.4: Characterization of the ASC-PYD subunit interaction.
(A) Secondary chemical shifts of monomeric ASC-PYD in solution and (B) multimeric ASC-PYD in the filament form. Regions identified as α-helices are indicated at the top of each panel with white rectangles. (C) Chemical shift differences between solution and solid state chemical shifts for the nuclei ^{15}N, $^{13}C\alpha$, $^{13}C\alpha$. Chemical shift variations larger than the mean value are marked with different colour codes on the secondary structure elements. Residues that belong to the interfaces of type I, II and III are marked in dark orange, dark yellow and dark red, respectively. (D) Structural location of the residues with significant chemical shift differences between the monomeric and filament form, as identified and using the same color code as in panel C. Amino acid side chains involved in intersubunit salt bridge interactions are shown as stick models and with their sequence label.

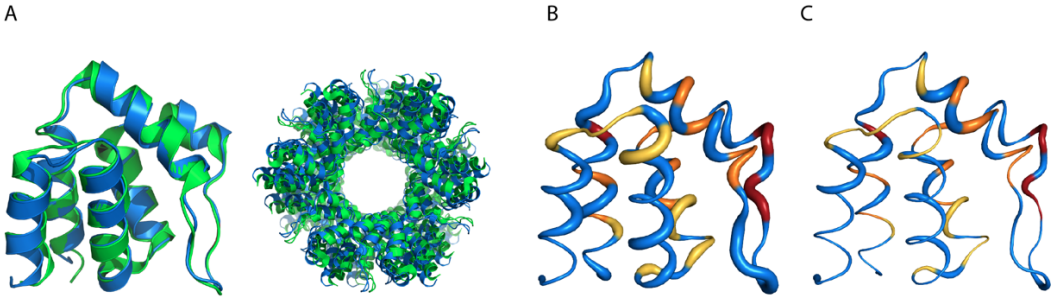


FIGURE 5.5: Structural comparison of mouse and human ASC filaments and evolutionary conservation. (A) Superimposition of ASC-PYD units within the assembled filament of mouse (blue) and human (green,PDB 3J63[116]) ASC and overlay of the full filament structure.(B) Evolutionary conservation of residues in the mammalian ASC-PYD. Structure of the mouse PYD where low and high conservation is indicated with thin or thick putty cartoon representation, respectively.The sequence alignment includes mouse ASC-PYD (Q9EPB4), human ASC-PYD (H3BP42), bovine ASC-PYD —Q8HXK9— and human PYDC1 (Q8WXC3).Residues defining the interface types I, II and III are coloured in dark orange, dark yellow and dark red, respectively.(C) Conserved residues of mouse ASC-PYD within selected mouse PYD domains from NLRs and PHYHIN receptors.The structure-based sequence alignment includes mouse ASC_PYD (Q9EPB4) AIM2_PYD (Q91VJ1), NALP3_PYD (Q8R4B8), NAL10_PYD (Q8CCN1). The sequence conservation analysis is performed as reported above

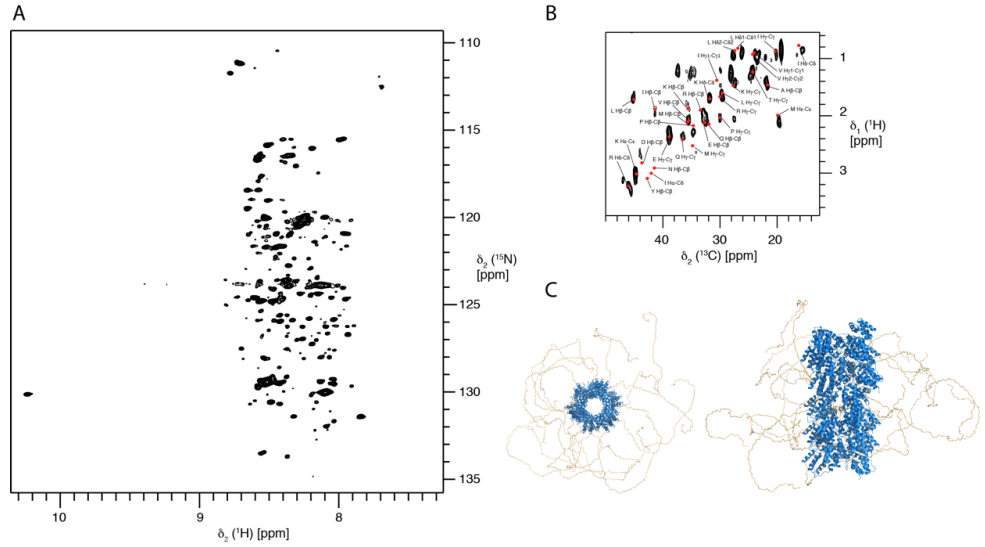


FIGURE 5.6: The CARD domain in the mouse ASC-FL filament is flexibly unfolded. (A) Solid state 2D [^{15}N,^{1}H]-HSQC spectrum of [*U*-^{15}N]-labelled ASC-FL. (B) Refocused carbon-detected INEPT spectrum of recombinant, [*U*-^{13}C]-labelled mouse ASC-FL. Random coil chemical shift positions for the 20 common amino acid are reported as red dots. (C) Structural model of the full-length ASC filament. Multiple ASC-PYD domains (blue) provide the structural scaffold for the filament formation while the ASC-CARD domains (yellow) remain unfolded and flexible relative to the ordered central filament core.

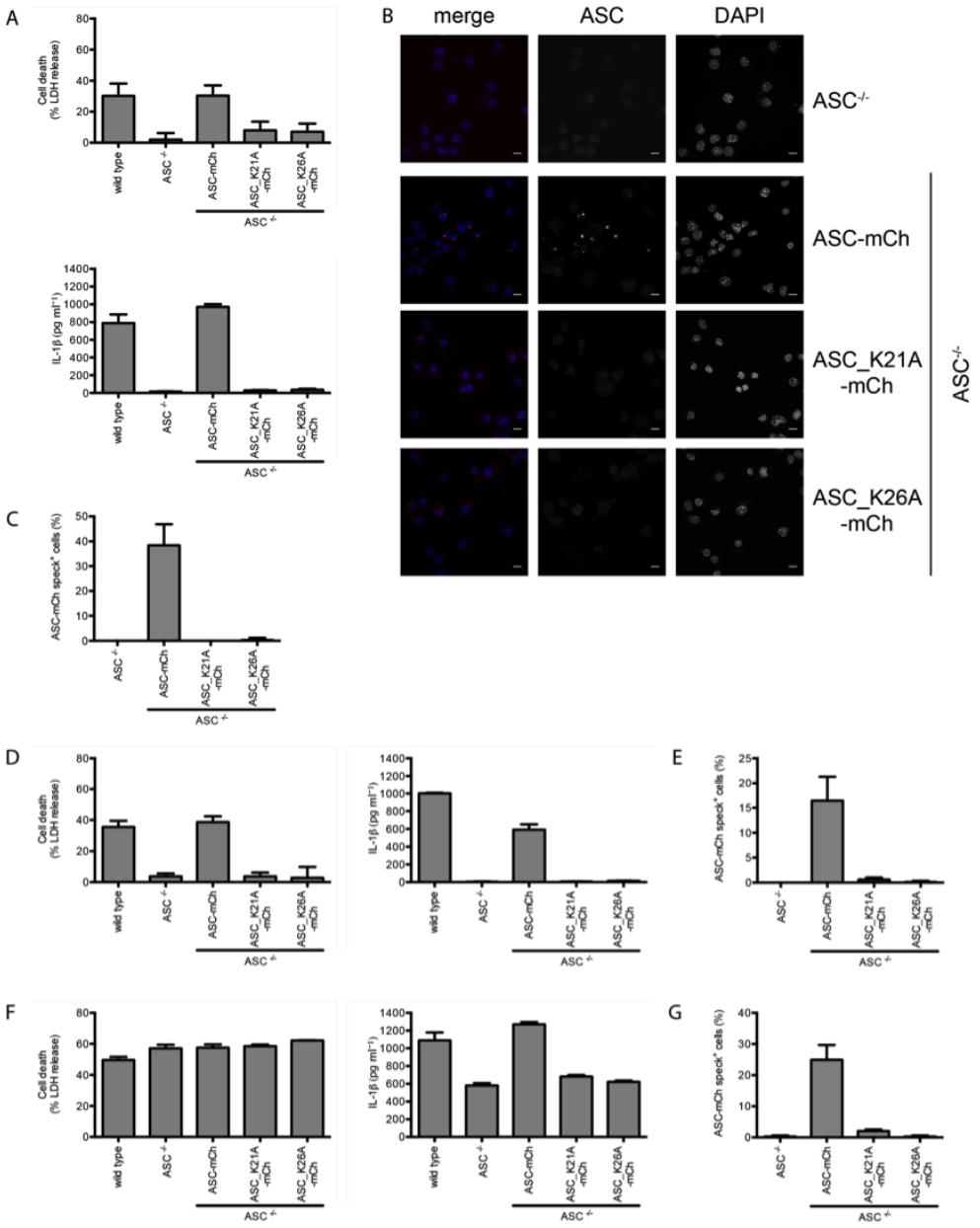


FIGURE 5.7: Effect of single point mutations on ASC-dependent signalling. Cell death as measured by LDH release, IL-1β secretion and ASC speck formation in LPS-primed immortalized mouse macrophages from the indicated genotypes: Wild-type (WT), $Asc^{-/-}$ or $Asc^{-/-}$ expressing endogenous levels of wild-type ASC-mCherry or ASC-mCherry K21A and K26A. Macrophages were stimulated with 5 mM ATP to activate NLRP3 (A-C), transfected with 1 μg/ml poly(dA:dT) to trigger AIM2 (D-E) or infected with log phase *S. Typhimurium* at an MOI of 10 to activate NLRC4 (F-G). Scale bars 10 μm.

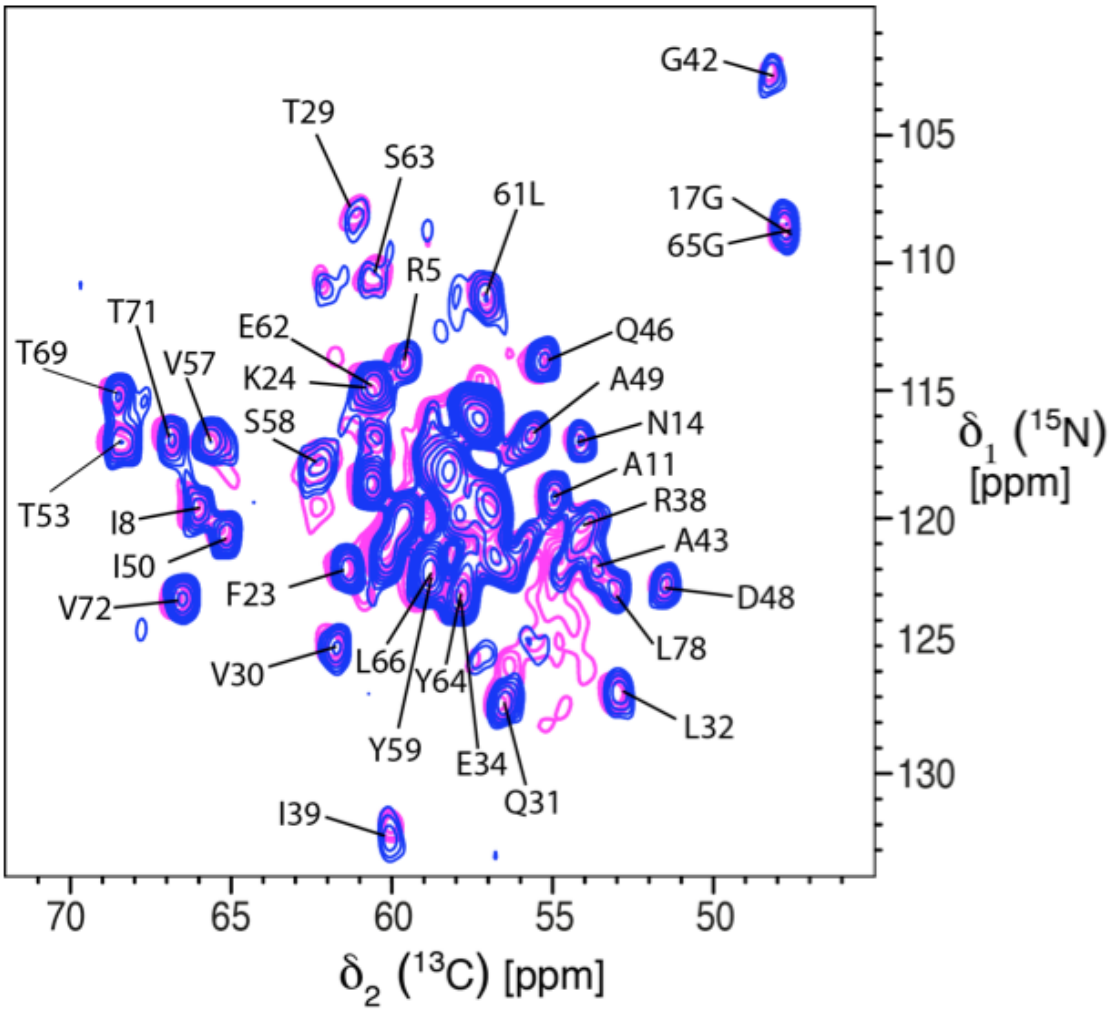


FIGURE 5.8: Sequence-specific NMR resonance assignments of ASC-PYD and ASC-FL filaments. Superimposition of solid-state 2D-NCA correlation spectra of the ASC-FL (magenta) and ASC-PYD (blue) fibrils. Sequence-specific assignment is shown for the less crowed region of the spectra using the one-letter amino acid code and residue number.

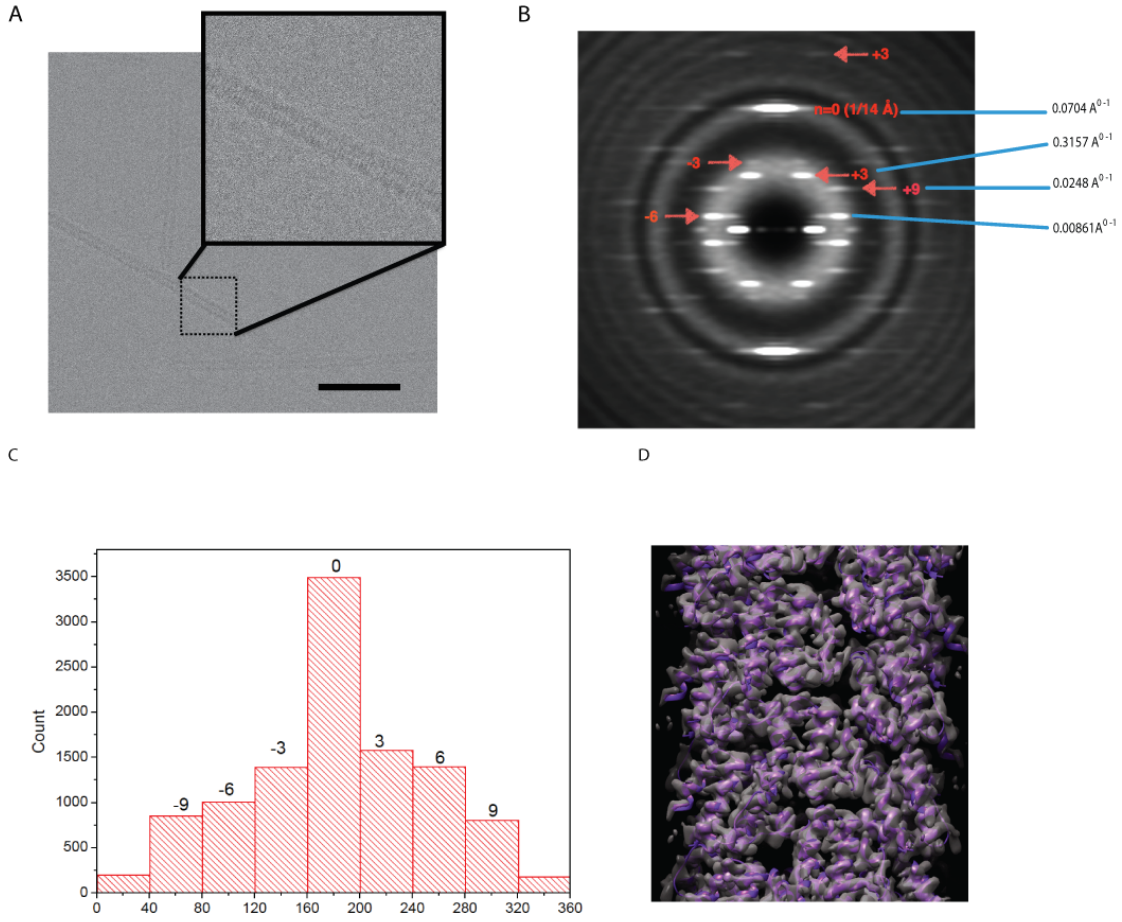


FIGURE 5.9: cryo-EM ASC-PYD structure calculation. (A)Higher magnification cryo-EM image of ASC-PYD filaments. Fibrillation of the PYD domain achieved by one step 50 times dilution of denatured protein with phosphate buffer + 300 mM NaCl and incubation for ∽20 minutes. The filaments are, well-ordered and separated filaments, which is requirement for structure determination by cryo-EM. (B) Layer line analysis of the average power septra obtained from ∼ 20000 of ASC-PYD filament segments, indicating a helical structure with an axial rise 14.2Å (n=0 layer line) per subunit in real space. Due to well-ordered nature of the filaments, there are no variable twist features. (C) The filament segments divided into separate bins according to their out of plane tilts, which ranged from -12^0 to 12^0. Scalebar:50nm

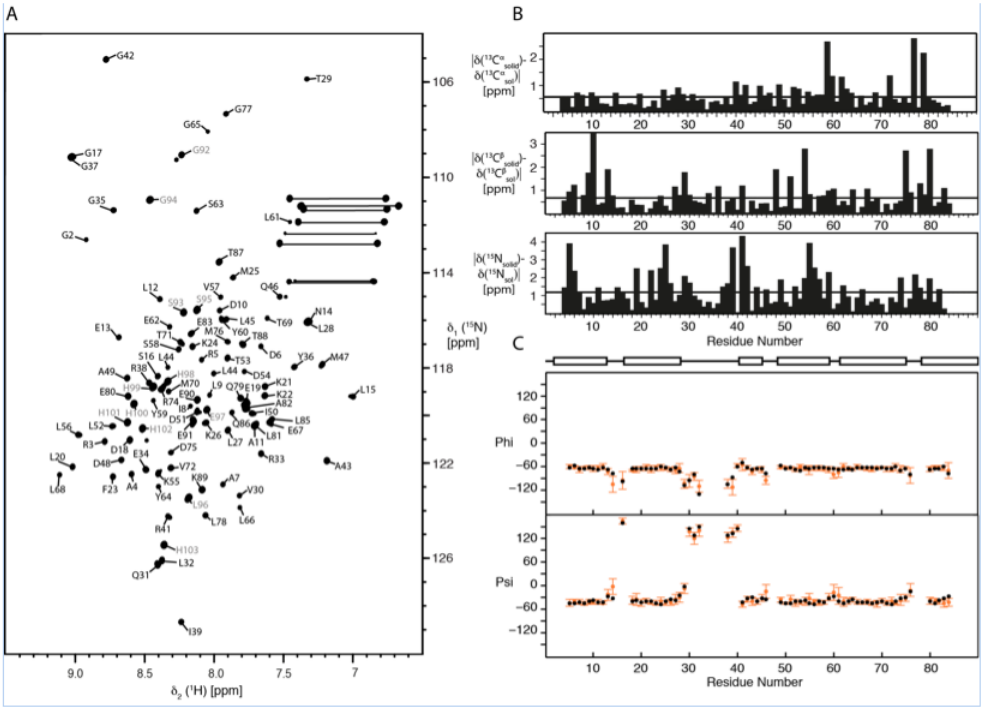


FIGURE 5.10: Sequence-specific resonance assignments of monomeric ASC-PYD and comparison with the ASC-PYD fibril form. (A) 2D [^{1}5N- ^{1}H]-HSQC spectrum of monomeric [U-^{1}5N]-labelled ASC-PYD recorded in 50 mM glycine buffer and 150 mM NaCl, pH=3.7 at 25^{0}C. Sequence-specific resonance assignments obtained from 3D triple resonance solution NMR experiments is shown using the one-letter amino acid code and residue number. Correlation of NH_2 moieties of Gln and Asn side chain are connected with solid lines. Cross-peaks that belong to the C-terminal tag are labelled in gray. (B) Comparison of isotropic chemical shift of monomeric ASC-PYD in solution versus oligomerized solid state form computed as the absolute values of the differences (Solid–Solution). The mean values for 13Cα (0.56 ppm), 13Cβ (0.68 ppm) and ^{1}5N (1.18 ppm) are indicated as a black line. (C) Comparison of TALOS+ predicted backbone dihedral angles for monomeric ASC-PYD in solution (black dots) versus oligomerized solid state form (orange dots). Comparison is performed for those residues with statistically significant predictions and error bars for solid state values are based on the 10 best database matches.

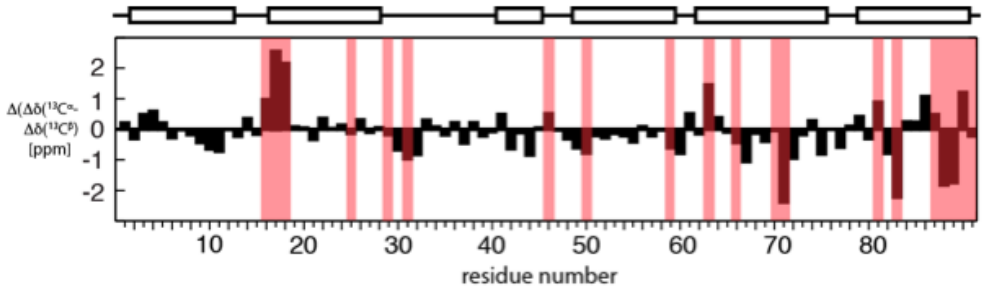


FIGURE 5.11: Comparison of secondary structure of the mouse and human monomeric ASC-PYD. (A) Secondary backbone ^{1}3C chemical shift differences versus residue number of the monomeric ASC-PYD from mouse and human (BMRB accession number 15313). Random coil backbone conformations have a value of zero in this representation. Non-conserved residues between mouse and human ASC-PYD sequences are marked with red bars. Protein regions identified as α-helices are indicated at the top of the panel.

TABLE 5.1: Supplementary Table 1.
Structural statistics for the mouse ASC PYD filament

	Protein
NMR distance and dihedral constraints	
Distance constraints	
Total	674
Intra-residue	19
Inter-residue	655
Sequential ($\|i–j\| = 1$)	89
Medium-range ($\|i–j\| < 4$)	209
Long-range ($\|i–j\| > 5$)	179
Intermolecular	65
Hydrogen bonds	0
Total dihedral angle restraints	140
φ	70
ψ	70
cryo-EM	
Resolution(Å)	3.7
No. reflections	50317 (work: 48264, test: 2572)
Rwork / Rfree	0.32/0.33
Structure statistics	
Violations (mean and s.d.)	
Distance constraints (Å)	0
Dihedral angle constraints (°)	2(0)
Max. dihedral angle violation (°)	10.5(0.3)
Max. distance constraint violation (Å)	0.0
Deviations from idealized geometry Bond lengths (Å)	0.0028
Bond angles (°)	0.51
Impropers (°)	0.46
Ramachandran analysis Most favored regions	99.0%
Disallowed regions	0.0% 0.46
Average pairwise r.m.s. deviation** (Å)	
Heavy	0.48
Backbone	0.14

**Pairwise r.m.s. deviation was calculated among 10 refined structures.

Chapter 6

6. Exploring the interactome: Microfluidic isolation of proteins and interacting partners for visual analysis by quantitative electron microscopy

Dominic Giss,[1,2]Simon Kemmerling,[1,2], Venkata P. Dandey[1,2],Henning Stahlberg[1,2], Thomas Braun[1,2*]

The following section has been published in:

Analytical Chemistry 86(10), 4680-7 (2014) (doi: 10.1021/ac4027803.)

1. *Center for Cellular Imaging and NanoAnalytics (C-CINA), Biozentrum,*
2. *Biozentrum, University of Basel, Klingelbergstrasse 70, 4056 Basel, Switzerland*

*Corresponding authors:Thomas.braun@unibas.ch

My contribution to this study was to implement the semi automate EM tool chain for data collection of negatively stained sample and its data processing for quantitative analysis at single molecule level.

Abstract

Multimolecular protein complexes are important for many cellular processes. However, the stochastic nature of the cellular interactome makes the experimental detection of complex protein assemblies difficult, and quantitative analysis at the single molecule level essential. Here, we present a fast and simple microfluidic method for (i) the quantitative isolation of endogenous levels of untagged protein complexes from minute volumes of cell lysates under close to physiological conditions and (ii) the labeling of specific components constituting these complexes. The method presented uses specific antibodies that are conjugated via a photocleavable linker to magnetic beads that are trapped in microcapillaries to immobilize the target proteins. Proteins are released by photocleavage, eluted and subsequently analyzed by quantitative transmission electron microscopy at the single molecule level. Additionally, before photocleavage, immunogold can be employed to label proteins that interact with the primary target protein. Thus, the presented method provides a new way to study the interactome and, in combination with single molecule transmission electron microscopy, to structurally characterize the large, dynamic, heterogeneous multimolecular protein complexes formed.

6.1 Introduction

Systems biology aims to identify and quantify the molecular elements of biological systems, to determine the dynamic interactions between them, and to integrate the resulting information into system networks [130]. The development of these holistic models necessitates the quantitative detection of protein-interaction assemblies and their characterization in terms of stoichiometry, structure and temporal persistence. The stochastic nature of biological processes makes the analysis of these heterogeneous, multimolecular complexes at the single molecule level essential. Many experimental strategies are used today to characterize the proteome of cells [131],[132] or to define the interactome.[133],[134] In these, protein complexes are often identified using hybrid approaches relying on affinity purification or immunoprecipitation in combination with chemical cross-linking[135] or gel electrophoresis[136] and mass spectrometry[137]. These methods analyze a sample from multiple cells, and quantify the average relative occurrence and composition of the available protein complexes. Due to the large number of protein copies required, these methods are generally not well suited for the analysis of

heterogeneous, multimolecular protein complexes, such as the human Mediator complex [138], transiently formed during interaction pathways. Further, although proteins interacting within a complex can be defined, [139] the structure of the entire multimolecular complex remains unknown. Yet, the structure of the entire complex has to be known to fully understand the role of the complex in the cell. Electron microscopy can be used to image single molecules and obtain various types of structural information [140, 141]. For example, cryo-electron tomography (cryo-ET) can provide information at molecular resolution for larger proteins in unperturbed single cells [142, 143]. However, visual proteomics by electron tomography [144, 145] is restricted to small cells or to sections of cells (the sample diameter must be below 1 μm for cryo-ET to reach higher resolution) [146] , and protein recognition is limited to relatively large protein complexes [147, 148]. Further, the identification of complexes and their constituents without electron dense labels [149] in intact cells is challenging. We have recently developed a new approach to visual proteomics in which eukaryotic cells are lysed and the proteins in the lysate and membrane fragments visually analyzed by single molecule transmission electron microscopy (TEM) [150, 151]. Again, without the aid of electron dense labels interpretation can be ambiguous. However, in this case the cell lysate is easily accessible for labeling methods. Here we present a fast microfluidic protein isolation method and demonstrate its use with TEM for the extraction of protein assemblies from cell lysate and their quantification. The method allows endogenous levels of untagged protein complexes to be specifically isolated from minute volumes of complex biological background, and investigated at the single molecule level. It provides initial structural information. Importantly, the approach also allows interacting partners to be identified by a labeling step, here termed interaction-labeling, and the intact complex to be visualized by TEM.

6.2 Experimental section

6.2.1 Working principle

The approach is based on antibodies (ABs) that are conjugated to a photocleavable biotin crosslinker [152] (Figure 6.1). These conjugates interact with streptavidin-coated magnetic beads that are immobilized in microcapillaries by a magnetic trap (Figure 6.1B; Supporting Information Figure 6.7). The basic protocol is as follows: Cell lysate is mixed with biotinylated ABs against the target protein and passed over the trapped beads. Subsequently, the capillary is rinsed with

buffer to wash out all unbound proteins. Finally, the immobilized target proteins are released by illuminating the capillary with UV-light at a wavelength of 365 nm to photocleave the crosslinker. Target proteins are then eluted and loaded on TEM grids (Figure 6.1C). This protocol can be extended by two variations before photocleavage and elution: (i) A labeling step using immunogold can be included to identify and localize the target proteins or their copurified interacting partners in a protein assembly; (ii) extracted proteins can be used as bait to 'fish' for potential binding partners, again followed by a labeling step.

6.2.2 Separation setup

A 250 μm inner-diameter fused silica capillary (BGB Analitik AG #TSP-250350, Switzerland) was guided through a lens tube construction cube (Thorlabs #SM1C6, Germany) so that optical components could be easily mounted, and to provide protection from UV-light (Supporting Information Figure 6.7). The outer polymer coating of the capillary was thermally removed from a length of 1 cm-long region at the center of the cube to avoid absorbance of UV-light. Magnetic beads were trapped in the stripped region by two external permanent magnets (Supermagnete #Q-20-10-05-N, Switzerland) positioned below the capillary at the center of the cube. UV-light at 365 nm was emitted by a high power UV LED (Thorlabs #M365L2, Germany) mounted on top of the cube and guided onto the magnetic beads using collimator and focusing lenses (Thorlabs #LA1952 and #LA1131, Germany). Liquids were aspirated or dispensed via a stepper motor syringe pump (Genie Pump, Kent Scientific) from 200 μl PCR tubes mounted on an xy-stage.

6.2.3 Loading of magnetic beads

About 40*106 streptavidin-coated magnetic beads with a diameter of 1 μm (corresponding to 4 μl of the beads employed; Dynabeads® MyOne^{TM}, Invitrogen #656-01, Switzerland) presented in phosphate buffered saline (PBS) were aspirated into a 250 μm inner-diameter fused silica capillary at a flow rate of 20 μl/min. An external permanent magnet positioned in close proximity below the capillary caused a plug of beads to form (Supporting Information Figure 6.7). After this bead-loading step, a second magnet was added to fully trap the plug, and the beads were rinsed with the buffer used for the subsequent experiment. In this manner, a robust, 3 - 4 mm long plug was formed and equilibrated with the required buffer.

6.2.4 Antibody biotinylation

ABs were incubated with a 10-fold molar excess of photocleavable NHS-Biotin crosslinker (Ambergen, USA) for 1.5 h at pH 8.2. After incubation, the mixture was dialyzed overnight against 2 l of PBS (phosphate buffered saline, 0.01 M phosphate buffer 0.0027 M potassium chloride and 0.137 M sodium chloride, pH 7.4, Sigma-Aldrich #P4417, Switzerland) using dialysis buttons (membranes with a 13 kDa cut-off) to remove unbound crosslinker.

6.2.5 Extraction and purification of target structures

A stock solution of apoferritin (AF) from equine spleen in 50% glycerol (Sigma-Aldrich #A3660, Switzerland) was diluted to 3 - 300 nM using baby hamster kidney (BHK) cell lysate. The AF / cell lysate mixtures were incubated for one hour on ice with biotinylated polyclonal anti-horse spleen ferritin ABs produced in rabbit (Sigma-Aldrich #F6136, Switzerland; henceforth termed anti-ferritin ABs) at 47 nM concentration. Subsequently, 2 μl of the mixture was passed over trapped magnetic beads for 15 min to immobilize target proteins on the beads. The actual incubation time of the biotinylated ABs with the streptavidin-coated beads was about 60s on average. Finally, the capillary was rinsed with PBS to remove unbound proteins. Progress of the washing process was probed by collecting a series of eluted wash fractions and analyzing them by TEM. To demonstrate that immobilized proteins can be used as bait for potential binding partners, 2 μl of a 1 nM solution of polyclonal anti-horse ferritin antibodies produced in goat (Lubio Science #A70-128A, Switzerland), were diluted with BHK cell lysate and passed over previously immobilized AF for 15 min. After incubation, the capillary was rinsed with PBS to completely separate trapped ABs from cell lysate. Immobilized goat anti-ferritin ABs were then labeled with 10 nm colloidal immunogold by rinsing the capillary with 2 μl of anti-goat IgG gold (Sigma-Aldrich #G5402, Switzerland) diluted 1:10 in 0.1% BSA (Sigma-Aldrich #A3294, Switzerland), for 15 min. To extract endogenous 26S proteasome, HEK293 cell lysate was incubated with biotinylated polyclonal anti-20S proteasome ABs from rabbit (Merck Millipore #ST1053) for 120 min on ice. 4 μl of the mixture, corresponding to the lysate of about forty thousand cells, were then passed over immobilized magnetic beads during a time period of 15 min. Washing steps were carried out with HEPES buffer (25 mM HEPES-KOH, pH 7.4; 5 mM MgCl2 and 10% glycerol). The isolation of endogenous TCP-1 ring complex (TRiC) from HEK293 cell lysate

was performed using biotinylated monoclonal anti-CCT epsilon ABs (Pierce Antibodies #MA1-82643) and HEPES buffer (20 mM HEPES-KOH, pH 7.4; 5 mM MgCl2; 50 mM NaCl and 10% glycerol).

6.2.6 Recovery of target structures

Immobilized target proteins were released from the magnetic beads by illuminating the plug in the capillary with UV-light for 10 min to cleave photocleavable biotin. Photocleavable biotin shows maximal absorbance around 270 nm, but still has distinct absorbance at 365 nm, the wavelength emitted by the UV LED employed. After cleavage, analytes were eluted in 3 - 6 μl of buffer and collected.

6.2.7 Washing procedure

After each incubation step, the capillary was rinsed with buffer (PBS or HEPES buffer, depending on the experiment) to separate unbound from immobilized proteins. Depending on the applied cell lysate concentration, the immobilized beads were washed for 5 - 25 min at 20 - 40μl/min flow rates. After an experiment, the magnetic beads were discarded by pressurizing the capillary. The latter was then extensively washed with double-distilled H20, 70% EtOH and 0.1 M NaOH.

6.2.8 Cell preparation

Adherent baby hamster kidney fibroblasts (BHK21; ECACC 85011433) and 293 cell line human embryonic kidney cells (HEK293; ECACC 85120602) were cultured in polystyrene T75-flasks containing 30 ml DHI-5 medium (see below) for BHK cells and 10 ml culture medium (see below) for HEK293 cells, at 37^0C and 5% carbon dioxide. To split the cells, the medium was removed and the flask was washed with 10 ml of 37^0C warm PBS w/o calcium and magnesium (Dulbecco's phosphate buffered saline, Sigma-Aldrich #D8537, Switzerland). To detach the cells, 3 ml of trypsin-EDTA solution (0.05% trypsin 0.53 mM EDTA, Invitrogen #25300-054, Switzerland) were added and drained before the cells were incubated at 37^0C for 5 min. The detached cells were diluted with 10 ml of 37^0C warm medium and homogenized using a pipette. For BHK cells, 0.5 ml of the homogenized cell suspension and 30 ml of fresh medium were returned to the flask for further cultivation. For HEK293 cells, 1 ml of the homogenized cell suspension and 9 ml of fresh medium were used. The rest of the cell suspension was washed

twice with PBS and concentrated to a final cell concentration of 1000 cells/μl for BHK cells or 10000 cells/μl for HEK cells using PBS or HEPES buffer (depending on the experiment). In both cases, cells in 0.5 ml of the suspension were then physically lysed by sonication for 30 - 60s using a tip sonicator while cooling. The fresh lysates were either directly used for experiments or aliquoted and stored at -80^{0}C. DHI-5 medium is a 1:1:2 mixture of DME (Dulbecco's modified Eagles medium, Sigma-Aldrich #D6171, Switzerland), HamF12 (Nutrient mixture F-12Ham, Sigma-Aldrich #N8641, Switzerland), and IMDM (Iscove's modified Dulbecco's medium, Sigma-Aldrich #I3390, Switzerland) media, supplemented with 5% FCS (Fetal calf serum, Sigma-Aldrich #E7524, Switzerland) and complemented with non essential amino acids (MEM non-essential amino acid solution, Sigma-Aldrich #M7145, Switzerland), L -glutamine (L -glutamine solution, Sigma-Aldrich #G7513, Switzerland), and vitamins (RPMI1640 vitamins solution, Sigma-Aldrich #R7256, Switzerland). For HEK293 cells the culture medium was EMEM (Eagle's minimal essential medium, Sigma-Aldrich #M2279, Switzerland) supplemented with 10% FCS and complemented with nonessential amino acids and L–glutamine.

6.2.9 TEM grid preparation

Immediately after elution, 3 μl of the eluted analytes were incubated for 90 s on glow-discharged carbon coated copper TEM grids (200 mesh). After a blotting step, grids were washed five times for 12 s on double-distilled H20 and negatively stained on two 4.5 μl drops of 2% uranylacetate. After every incubation and the final washing step, excess liquid was removed using blotting paper.

6.2.10 Image acquisition and processing

Automated image acquisition was done on a FEI T12 operated at 100 kV using the Leginon 2.1 (incl. in Myami 2.1) software [32]. The images were recorded on a Gatan 2k x 2k CCD camera. A semi-automatic image acquisition routine was used for the quantitative analysis of analytes (Figure 6.2): First, five squares of the TEM grid were manually selected for analysis. Then, a mesh consisting of 35 sub-squares was created within every square, and 11 images per sub-square were automatically acquired at 12000x magnification (pixelsize: 0.8725 nm, defocus: 0.8 - 1.4 μm, dose: 22 - 35 electrons/$\mathring{A}^2$). Thereby about 12% of the area of a particular square was imaged. Subsequently, using the 1925 images collected per

grid, target proteins were identified by their visual appearance in negative stain, and counted. This was done manually and/or using the Appion 2.1 (incl. in Myami 2.1) software [153] in combination with a template picking routine [100]. Note, that ABs often linked targets to form multi-protein complexes. In this case, every target unit was counted individually. Manual TEM was carried out on a Philips CM10 operated at 80 kV. The images were recorded on a 2k x 2k CCD camera (Olympus SIS, Münster, Germany). For 2D class averages, images were acquired at 130000x nominal magnification (pixelsize: 0.37 nm, defocus: 0.1 – 0.3 μm. Particles were picked manually using EMAN2 software package [33], classified and averaged using the e2refine2d algorithm.

6.2.11 Results and discussion

We have developed a quantitative method that enables the fast and specific extraction and recovery of target proteins from minute volumes of cell lysate without the need for genetic engineering to introduce affinity tags. Samples are deposited on TEM grids and examined at the single molecule level by quantitative TEM (qTEM), which also delivers structural information by single particle analysis. Our method also allows potential binding partners to be detected and analyzed. These interaction partners either copurify with the primary target protein or can be "fished" from complex samples using a previously isolated protein as bait. Immunogold labeling both aids protein identification and localizes specific proteins in the extracted multimolecular complexes. To quantitatively assess the method, we complemented the crude cell lysate obtained from sonicated BHK cells with known amounts of AF as target protein and incubated this lysate with biotinylated anti-ferritin ABs. 2 μl of the mixture (total amount of AF, 6 - 600 fmol, depending on the experiment) were passed over

previously immobilized magnetic beads during a time period of 15 min, followed by 200 μl of PBS during a time period of 10 min. Photolysis was carried out for 10 min after this washing step, and the released AF-AB complexes were eluted in 6 μl of PBS for TEM analysis. The excess of streptavidin-coated magnetic beads offered ample binding surface for the biotinylated ABs; as expected control experiments showed that using more beads did not improve extraction efficiencies. An example of the isolation process is shown in Figure 6.3 . The first wash fractions after the immobilization step contained BHK cell lysate and some non-extracted AF (Figure 6.3B), recognizable by its typical ring-shaped projection (compare Supporting Information 6.8). Most of the unbound proteins were removed by

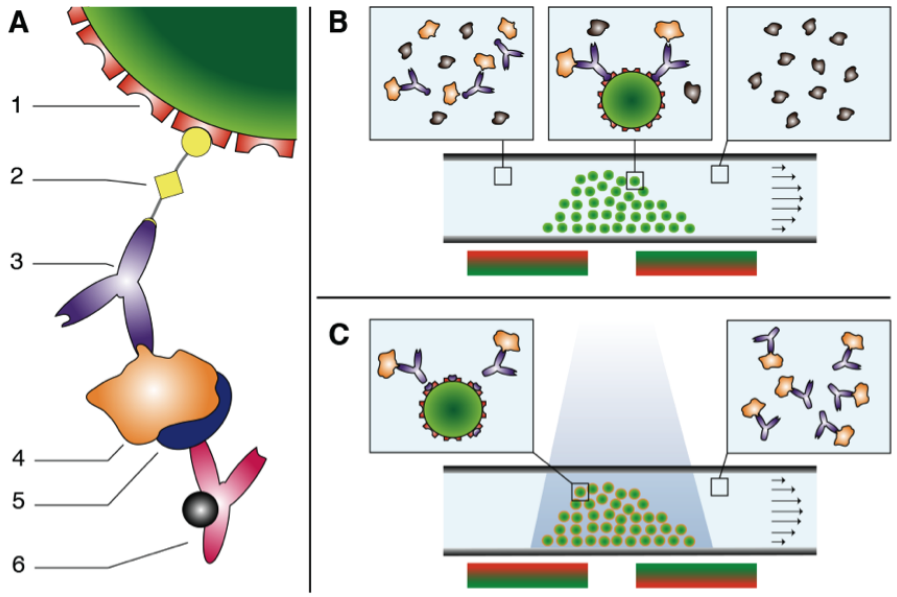


FIGURE 6.1: Principle of the microfluidic affinity isolation method. (A) AB-biotin conjugates bound to a streptavidin-coated magnetic bead. 1: Streptavidin-coated magnetic bead of 1 μm in diameter; 2: photocleavabcle NHS-Biotin crosslinker; 3: AB against the target protein; 4: target protein; 5: protein binding partner (optional); 6: immuno-gold labeled AB against the binding partner. (B) Affinity extraction. Streptavidin-coated magnetic beads are trapped in a microcapillary (inner diameter of 250 μm) using two external magnets positioned underneath the capillary. After loading, the capillary is rinsed with sample solution that was previously incubated with biotinylated ABs against the target structure. The capillary and beads are then washed to remove other proteins; ABs and their targets remain immobilized. Optionally, the capillary can next be rinsed with potential binding partners of the target structures and/or with immuno-gold ABs to label immobilized target proteins or their interacting partners. (C) Photocleavage and elution. Target structures are recovered by illuminating the beads with UV-light at a wavelength of 365 nm and eluting the released AB-target protein complexes. Downstream applications such as TEM analysis follow.

the washing procedure (Figure 6.3C). AF-AB complexes were recovered after the subsequent photocleavage step (Figure 6.3D). The clean background in this image documents the very low concentration of contaminants on the TEM grids. Proteins that were nonspecifically bound to the beads were not released by the photocleavage step (compare Figure 6.3D and Supporting Information Figure 6.9). We performed qTEM to analyze the results of the isolation experiments. A semi-automatic, TEM image acquisition and data analysis procedure was developed for this (Figure 6.2). Thereby, 1925 images were recorded from defined regions of each TEM grid and the target proteins were detected by their visual appearance and counted, allowing the total amount of recovered protein to be measured. Running a series of experiments as described above with concentrations of the target protein AF ranging from 3 - 300 nM and an AB concentration of 47 nM, yielded the signal transfer function (SiTF, Figure 6.4and Supporting Information Figure 6.10) of the method for AF. This quantitative measurement

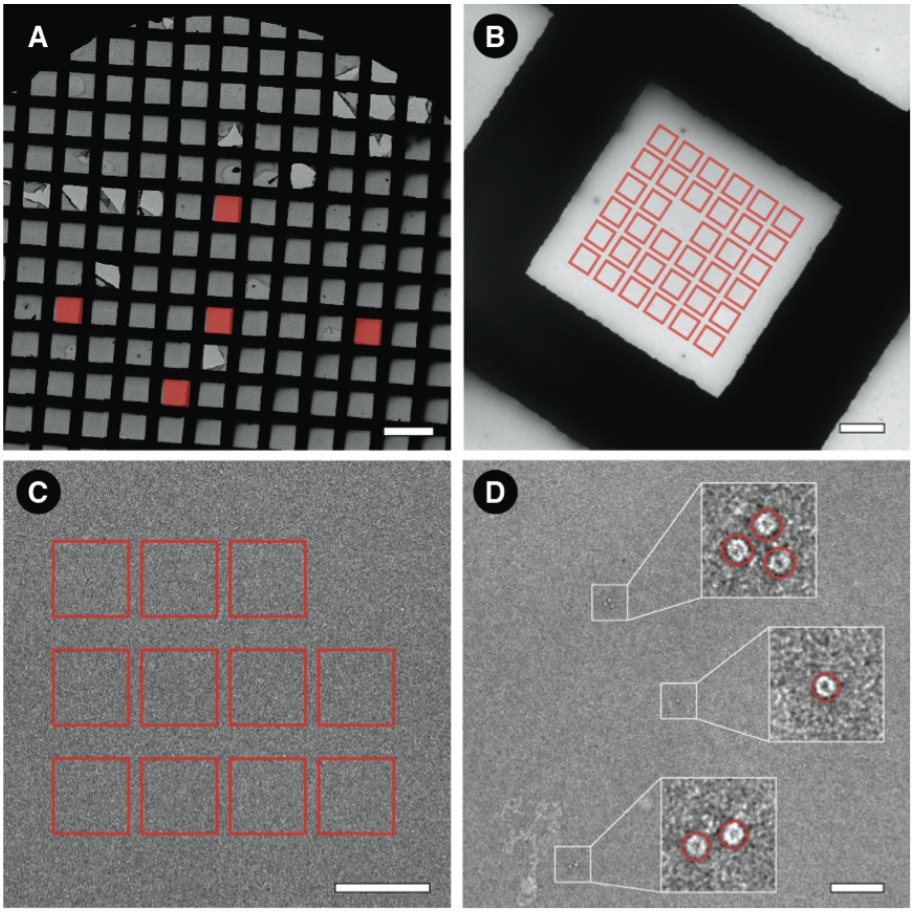


FIGURE 6.2: Semi-automatic TEM procedure for image acquisition and data-analysis. (A) Step 1: Manual selection of five squares of a TEM grid. (B) Step 2: The procedure targets 35 sub-squares per grid square. (C) Step 3: 11 images are recorded per sub-square. (D) Step 4: Image processing. Particles of interest are detected by their visual appearance in negative stain on the total of 1925 images per grid either manually or automatically using a template, and counted; each AF was counted as one particle, regardless of whether it was associated with other AFs or not. Scale bars, 200 μm (A), 20 μm (B), 2 μm (C) and 200 nm (D).

reflects the AB-antigen binding curve (Figure 6.4). Using the Hill equation for non-cooperative binding to extract the apparent dissociation constant Kd of the AB-AF interaction from the recorded data, gave an apparent Kd of 12 nM. Control experiments showed that the SiTF of AF to TEM grids depends linearly on the applied concentration of AF (Supporting Information Figure 6.11). These experiments show, that the AB binding characteristics dominates the SiTF of the presented isolation method.

We used HEK293 cell lysate to demonstrate the extraction and recovery of endogenous levels of protein complexes from cell extract by our one-step isolation method. In these experiments, we added biotinylated anti-20S proteasome ABs to the cell lysate to bind endogenous 26S proteasomes and processed the lysate of about forty

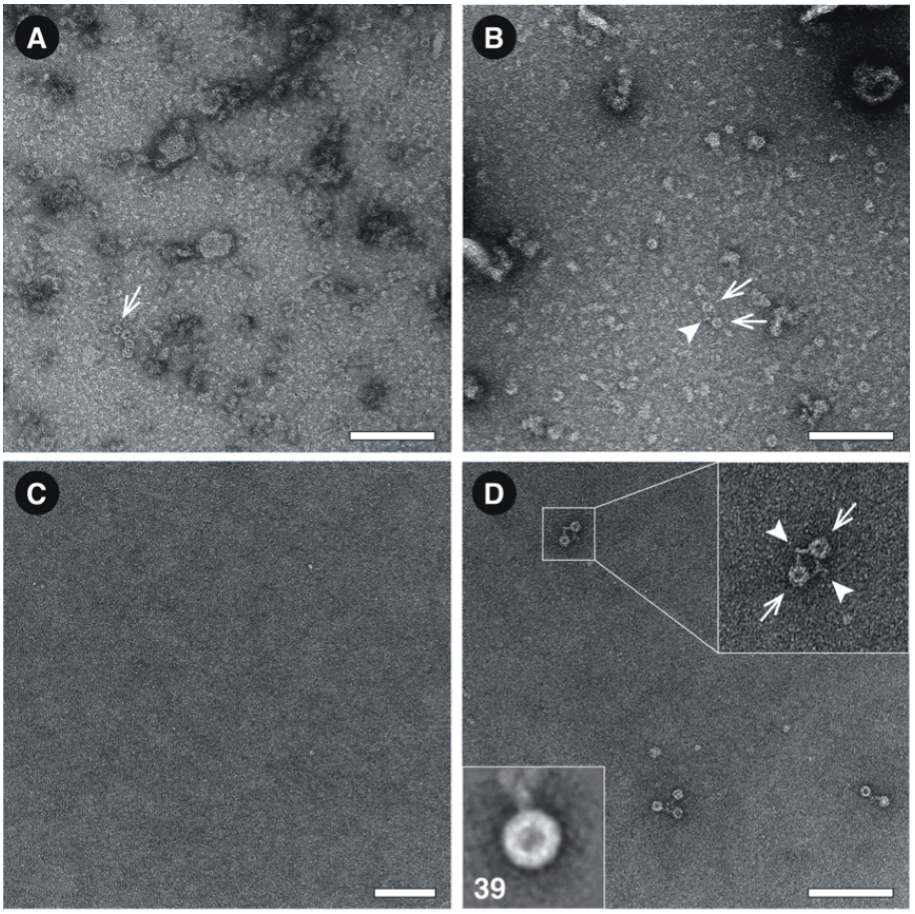


FIGURE 6.3: Affinity extraction and recovery of apoferritin. (A) The initial cell lysate / 30 nM AF (arrow) mixture. The mixture was incubated with biotinylated anti-ferritin ABs (47 nM) and passed over streptavidin-coated magnetic beads immobilized by a magnetic trap in a microcapillary, and the capillary was subsequently washed with buffer. (B) First wash fractions document that crude cell lysate does not obstruct the capillary and that not every single AF (arrow) – AB (arrowhead) complex was extracted. (C) Last wash fraction, confirming that unbound proteins were removed by the washing process. (D) Eluate after photocleavage of the crosslinker showing recovered AF (arrow) – AB (arrowhead) complexes; these have been cleanly isolated from the cell lysate. Inset: 2D class average (bottom left, number of averaged particles) of recovered AF with visible AB attached. Scale bars, 100 nm.

thousand HEK293 cells suspended in HEPES buffer. The protocol was similar to that described for AF. A typical experiment, starting with the lysis of cells and ending with the isolated target protein on TEM grids, took approximately 160 min, including an AB-target protein incubation step of two hours. As depicted in Figure 6.5, the method enabled us to isolate intact 26S proteasomes comprised of a 20S core complex and two weakly interacting 19S regulatory units from cell extract, as well as 20S proteasomes lacking one or both of the 19S regulatory complexes. The isolation of fully assembled 26S proteasomes, which are sensitive to buffer conditions (Supporting Information Figure 6.12), clearly demonstrates that complexes formed by weakly interacting proteins can be isolated by the proposed

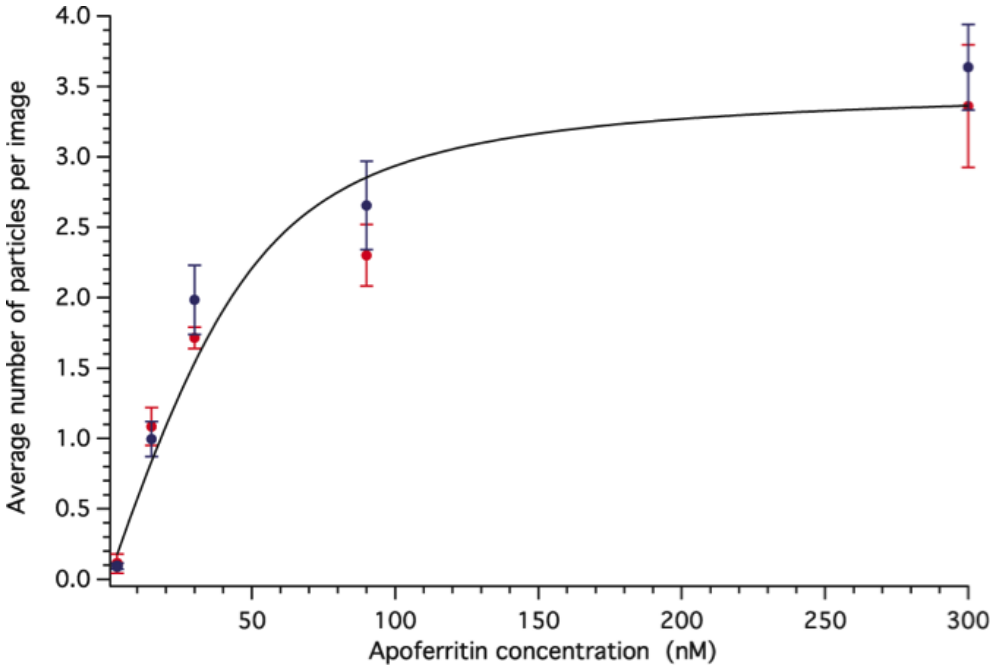


FIGURE 6.4: Signal transfer function of the microfluidic affinity isolation method for AF determined by qTEM. To quantitatively assess the response of the developed method, we processed different concentrations of AF, ranging from 3 - 300 nM, using a constant amount of anti-ferritin ABs (47 nM). Two TEM grids were prepared for every experiment (red and blue symbols). In a semi-automatic manner, 1925 images were collected per grid (385 per selected grid square), and AF was detected by its visual appearance in negative stain and counted (see Figure 6.2). The average number of AF particles per image (PPI) for the particular grid is shown. Error bars represent the standard deviation of the average PPIs of the five probed grid squares of the parent grid. The data reflects the AB-antigen binding curve for an increasing antigen concentration, which was modeled (black line) using Hill's equation.

method. Furthermore, using biotinylated anti-CCT ABs, we isolated endogenous levels of TRiC (Figure 6.5D). Note, that the application of unspecific ABs (or no ABs at all) did not result in the isolation of target proteins. The proteasomes and TRiC can be visually identified in negative stain TEM images due to their characteristic size and shape. The average structures calculated by single particle analysis are in excellent agreement with previously reported models [154–156] and 2D class averages of purified 20S proteasomes (Supporting Information Figure 6.13). Sometimes, ABs are visibly attached to the proteins, therefore we can be confident that the isolated particles are indeed the target protein complexes and that the corresponding structural analysis is relevant. Additionaly, immunogold markers can be used to unambiguously detect specific proteins and localize them within a complex. Exploiting this, we used immobilized target proteins to fish for potential binding partners and identified them by gold tags (Figure 6.1A). As a proof of concept, AF was mixed with BHK cell lysate and biotinylated rabbit anti-ferritin ABs and immobilized on magnetic beads as described above. To

mimic binding partners, the capillary was subsequently flushed for 15 min with a total of 2 μl of 1 nM goat anti-ferritin AB solution in BHK cell lysate. After washing, the capillary was rinsed for 15 min with 2 μl of anti-goat colloidal gold to label established protein complexes. Photolysis and the elution of AF-AB complexes for TEM analysis followed. As shown in Figure 6.6, the approach allows (i) binding partners (arrow) to be fished and recovered from complex backgrounds, and (ii) identified by an attached immunogold label (arrowhead). In our studies, we did not detect any cross-reactivity between anti-goat immunogold and rabbit ABs or AF. The presented method uses functionalized magnetic beads to form bead plugs in microcapillaries by magnetic trapping. This approach significantly reduces the diffusion path length required for protein binding and offers a massive increase in the available binding surface compared to, for instance, functionalized capillary walls. The technique provides a fast and convenient way to process of sub-microliter volume samples. Important for future high-throughput applications, the immobilization or release of beads can be easily tuned by adjusting the applied magnetic field gradient, enabling their rapid exchange [157].

The use of photocleavable linkers allows sample recovery under mild conditions that conserve the structural integrity of proteins. It is not necessary to use relatively harsh methods such as boiling the beads, treatment with low pH buffers, or changing the salt concentration, all of which might affect the structural integrity of target proteins or lead to disassembly of protein complexes. Indeed, the photocleavage employed is carried out at a wavelength of 365 nm, which is not significantly absorbed by most protein targets and therefore does not interfere with their structures or the complexes formed. Moreover, finite element simulations estimate that the UV-light induced temperature rise experienced by capillaries and buffers is lower than 0.6^0C (Supporting Information Figure 6.14. Elution by photocleavage is more specific than traditional chemical procedures and serves as a second "purification" step; proteins that are nonspecifically bound to the immobilized beads or the capillary surface are not released (Supporting Information Figure 6.9). This explains the low contaminant concentration in the eluates, as indicated by the almost perfectly clean background of the TEM images (Figure 6-3D, Figure 6.5 and Figure 6.6). The purity of the eluates facilitates qTEM and structural analysis of the individual complexes. Furthermore, we have demonstrated that weak interacting protein complexes, such as the regulatory 19S and the 20S core particle forming a 26S proteasome, survive the isolation procedure (Figure 6.6A&B). Quantitative TEM showed that the amount of recovered target protein depends on the initial concentration of the protein in the

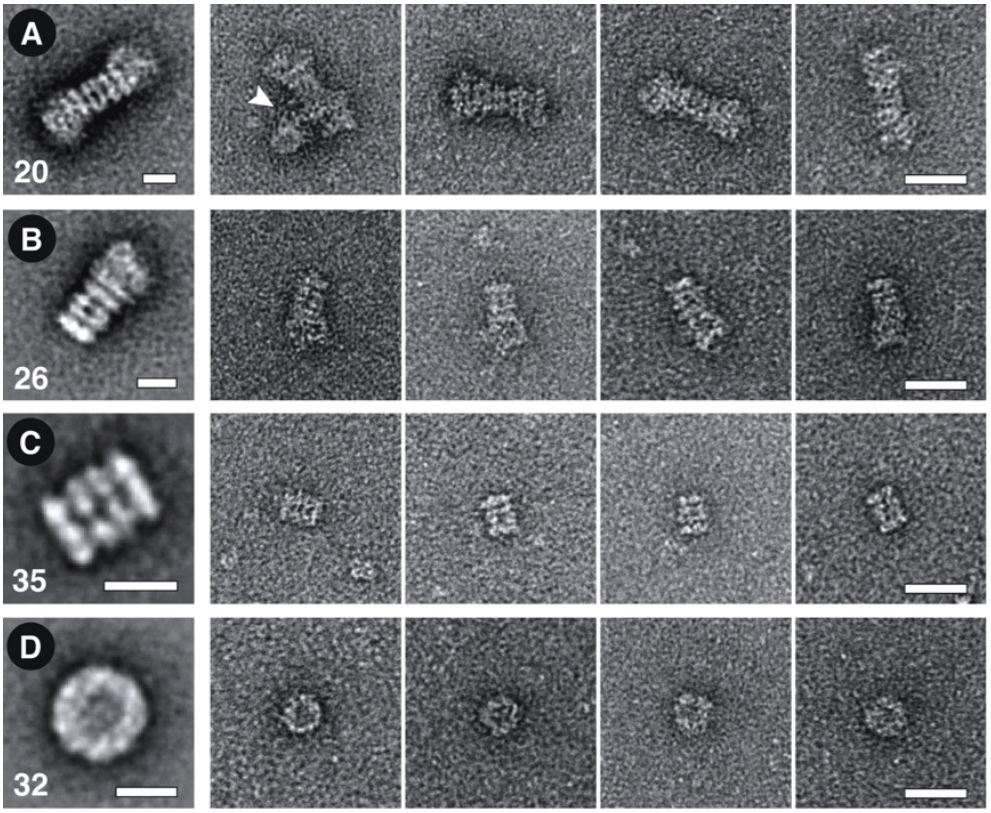


FIGURE 6.5: Isolation of endogenous protein complexes. HEK293 cell lysate was incubated with biotinylated anti-20S proteasome ABs (A, B, C) or biotinylated anti-CCT ABs, respectively (D). The lysate of about forty thousand cells was then passed over trapped streptavidin-coated magnetic beads to extract target complexes, which were subsequently recovered by photocleavage and transferred onto TEM grids. Left panels: Representative 2D class averages of isolated target structures (bottom left, number of averaged). Scale bar, 10 nm. Right panels: Gallery of negatively stained particles. Scale bar, 25 nm. (A) 26S proteasomes composed of two regulatory 19S particles and the 20S core proteasome. In rare cases, we found complexes linked by ABs (arrowhead). (B) Partially disassembled proteasomes, lacking one of the regulatory 19S particles. (C) Side view of isolated 20S core particles. (D) TRiC. Typical ring-shaped top views, 16 nm in diameter, are visible.

cell extract (Figure 6.4). Once the respective SiTF of a target protein has been calibrated (Figure 6.4 and Supporting Information Figure 6.10 and Figure 6.11), absolute quantities can be measured. To our knowledge, this is the first time where the feasibility of single molecule qTEM for protein quantification has been demonstrated. As shown in Figure 6.4, the SiTF of protein is dominated by the antibody-antigen binding curve, which also allows the apparent binding constant to be determined. Further, the method provides reproducible results; the resulting particle concentrations on two TEM grids prepared with the same sample differ by maximally 15% (peak-to-peak). The measurement variations of other methods used for protein quantification, e.g., reverse phase protein arrays, are comparable.

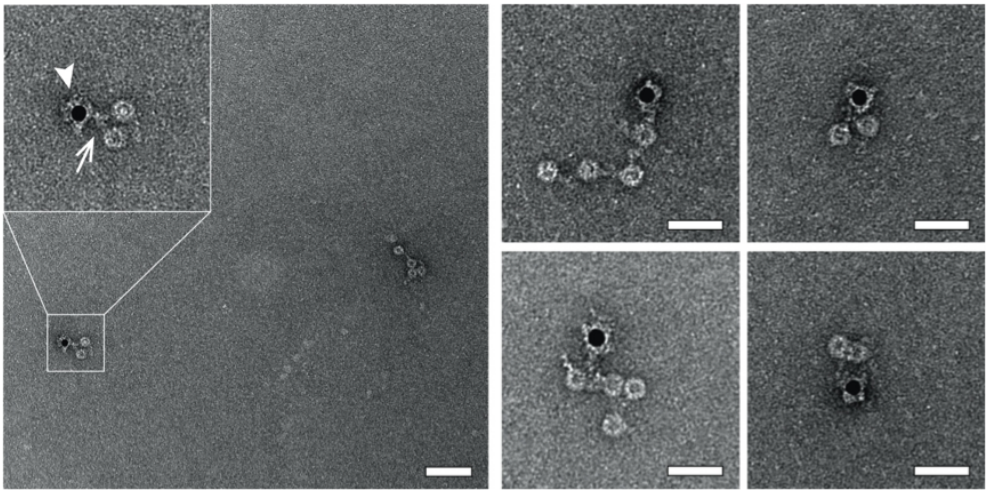


FIGURE 6.6: The detection of extracted protein binding partners by interaction-labeling. AF in a mixture of cell lysate and biotinylated anti-ferritin ABs was immobilized on the streptavidin coated magnetic beads, where it acted as bait protein. In a second step, the capillary was then rinsed with goat anti-ferritin ABs (arrow) supplied in cell lysate to mimic potential binding partners. In a third step, the capillary was rinsed with anti-goat colloidal gold (10 nm, arrowhead) to label the fished goat anti-ferritin ABs and, thus, facilitate the detection of established interactions by TEM. The complexes were recovered by photocleavage and transferred onto TEM grids. The black spots on the images are gold particles labeling the goat anti-ferritin ABs. Scale bars, 50 nm.

[158] As indicated in Supporting Information Figure 6.10, the SiTF of the presented isolation method depends on four main factors. Future improvements to their experimental aspects will result in more target proteins being isolated and detected: First, the use of smaller affinity molecules with only one binding site and higher binding constants, such as DARPins [159], nanobodies [160, 161] or aptamers [162] instead of ABs should increase the extraction efficiency of target structures. Second, the use of techniques enabling a much more efficient transfer of proteins onto TEM grids, such as the "sample writing" procedure that we described in Kemmerling et al., [151] will significantly increase the concentration of particles on TEM grids. Third, new camera types [126] will allow isolated particles to be imaged faster with a massively improved signal-to-noise ratio and counted more accurately. Fourth, improved algorithms for particle detection will allow the analysis of larger data sets with better precision. Together, this will make the image acquisition and/or quantitative analysis less time consuming and improve the sensitivity and detection limits of the technique. However, already now the method presented can be applied to collect enough images of well-distributed particles to calculate class averages of the isolated target structures (Figure 6.3 and Figure 6.5). The presence of ABs did not interfere with 2D class averaging (Supporting Information Figure 6.12). Class averages can aid protein identification

and provide initial structural information. Further, if their signal is not averaged out (inset Figure 6.3D), bound ABs revealed in the average can be used to localize specific subunits in a complex [156]. Figure 6.6 demonstrates the detection of extracted protein binding partners by immunogold labeling. In general, two problems must be overcome to detect binding partners by immunotagging: First, specific and efficient labeling must be achieved. Second, unbound label must be removed, or ways to discriminate between label specifically bound to the protein and label randomly adsorbed to the TEM grid must be found. The method presented provides an elegant and efficient way to remove unspecific label before elution of the protein complexes and TEM grid preparation, thus avoiding hurdle two.

6.3 Conclusions

We have presented a fast, one-step affinity isolation method for the quantitative extraction of endogenous levels of protein from complex samples without significant contaminants. Furthermore, protein-protein interactions can be probed and interacting partners labeled for subsequent structural analysis at the single molecule level by qTEM. This allows the composition, conformation and the structure of individual protein assemblies to be investigated. Moreover, quantitative conclusions can be drawn, since relative or even absolute quantification of protein levels can be achieved. This will allow the study of, e.g., pulse chase experiments inducing changes in protein complexes such as inflammasomes [163], not only on a functional and structural but also on a quantitative level. We foresee that in combination with novel grid preparation techniques, the use of single binding site affinity proteins and new class of detectors in electron microscopy, this methodology will facilitate the processing of single cell extracts by “lyse and spread” visual proteomics [150, 151, 164] and ultimately will offer a completely new way to study the interaction networks of protein complexes in individual cells.

6.4 Supporting Information

6.4.1 Photograph of the experimental setup

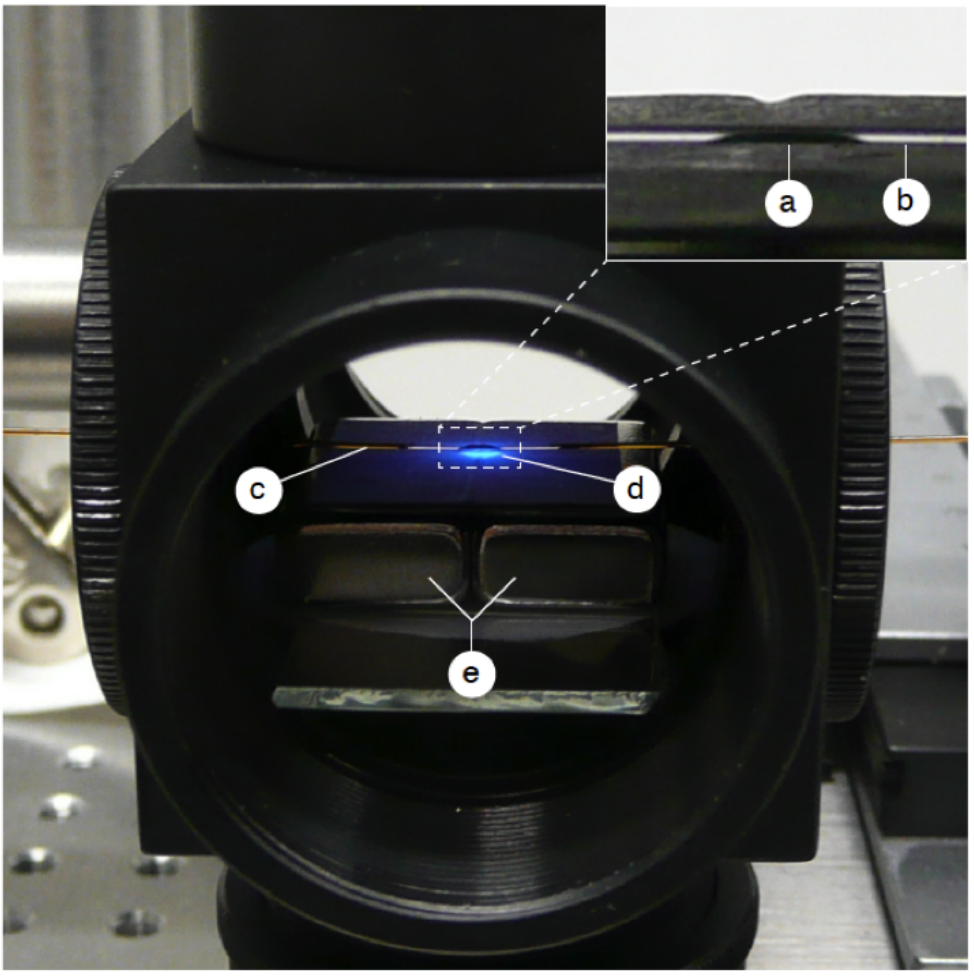


FIGURE 6.7: Photograph of the experimental setup and higher magnification photograph of a magnetic bead plug (inset). A fused silica capillary (c) with an inner diameter of 250 μm was guided through a lens tube construction cube. The outer polymer coating of the capillary was removed (b) over a length of 1 cm at the center of the cube to optimize the illumination conditions. About 40⋆106 magnetic beads (a) with a diameter of 1 μm were captured at the center of the cube by a magnetic trap consisting of two permanent magnets (e). Typically a bead plug with a length of 3 - 4 mm was created. UV-light with a wavelength of 365 nm emitted by a high power UV LED mounted on top of the cube was guided onto the capillary (d).

6.4.2 Electron micrograph and 2D class averages of apoferritin

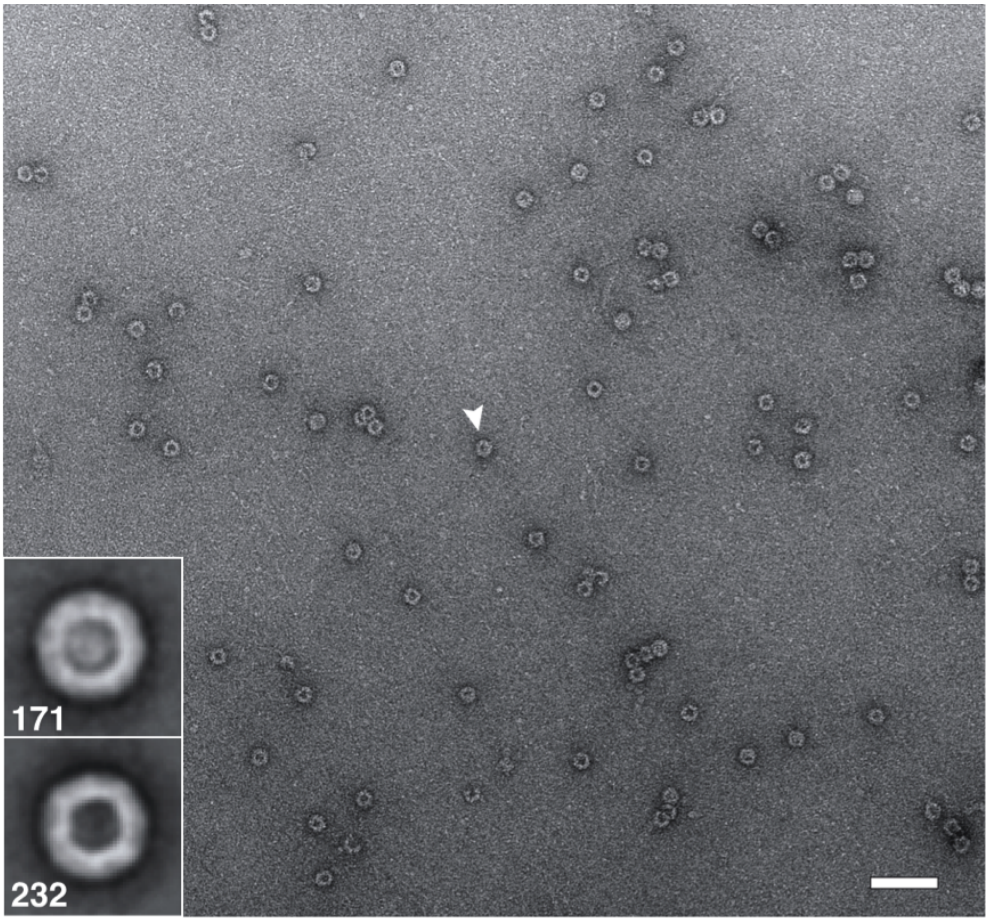


FIGURE 6.8: Negative stain transmission electron microscopy (TEM) of apoferritin (AF). A TEM grid was prepared applying a 200 nM solution of AF in PBS, negatively stained and imaged in the TEM. Ring-shaped projections (arrowhead) about 15 nm in diameter typical of AF (molecular weight, 443 kDa) are visible. Insets show representative 2D class averages (bottom left, number of averaged particles). Scale bar, 50 nm.

6.4.3 Electron micrograph and 2D class averages of apoferritin

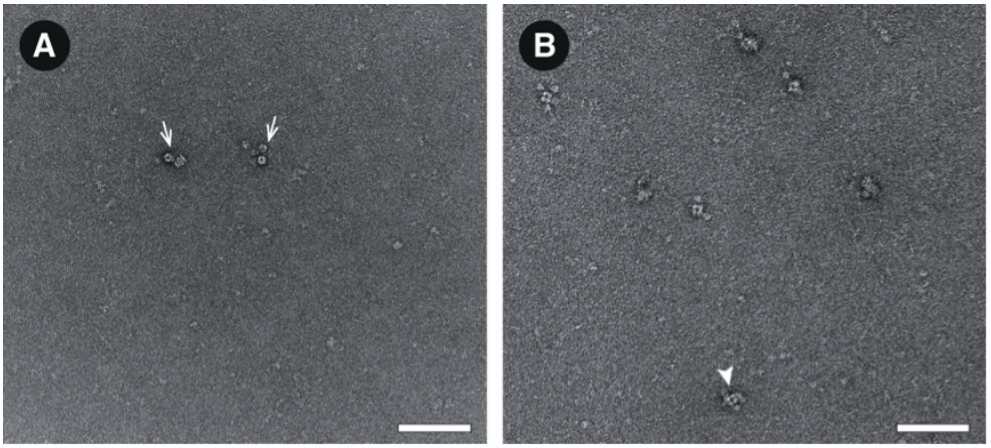


FIGURE 6.9: Negative stain electron micrograph of isolated 20S proteasomes (A) and nonspecifically bound protein (B). The 20S proteasomes of about one hundred thousand HEK293 cells were extracted and recovered using the presented approach. After photocleavage and elution of the released protein complexes (arrows, isolated 20S proteasome conjugates), the trapped beads were rinsed with 4μl of PBS complemented with 0.5 mM biotin (Sigma-Aldrich #B4501, Switzerland) to competitively release proteins that were nonspecifically bound to the beads or the fused silica surface of the capillary walls. After 5 min incubation, 6 μl of liquid was eluted and used to prepare TEM grids. The eluate contained a lot of protein, which must have been nonspecifically bound to the surfaces. In particular, we observed structures (arrowhead) that strongly resembled pyruvate dehydrogenase complexes [165] .

The results indicate that the elution by photocleavage acts as a second purification step, since such high concentrations of nonspecifically bound protein were never observed in the experiments. Scale bar, 100 nm.

6.4.4 Factors determining the signal transfer function of the method presented

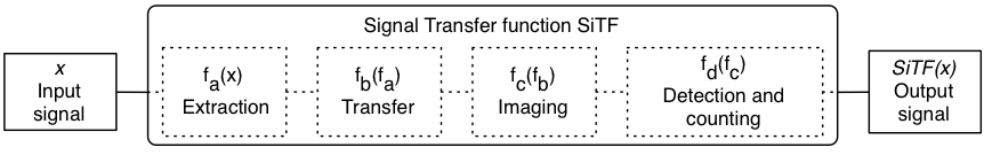


FIGURE 6.10: Factors determining the signal transfer function (SiTF) of the presented isolation method. The SiTF correlates the input signal x with the output signal SiTF(x). An example is presented in Figure 6.4, where the initial concentration of apoferritin (AF) particles (input signal) is correlated to the number of recovered AF particles (output signal). The individual factors influencing the SiTF are indicated in dotted boxes (fa-fd). fa: The extraction efficiency is mainly predetermined by the affinity of the antibody (AB) to the target protein, but also includes the binding efficiency of the ABs to the magnetic beads and the photocleavage efficiency. Note that the binding of ABs to the beads is expected to be very efficient due to the strong biotin-streptavidin interaction and the ample binding surface offered by the bead-plug. Furthermore, the AB concentration is a constant throughout the experiments. In general, fa(x) is a non linear function. If no extraction step is performed, fa(x)/x=1 and the SiTF is determined by the transfer, imaging and detection efficiency alone. fb: Transfer efficiency of the extracted protein to the transmission electron microscopy (TEM) grid. fc: Imaging efficiency, i.e., how well the imaging procedure documents the eluted sample (see Figure 6.2). fd: Detection and counting efficiency, i.e., how reliably the target protein is recognized and counted.

Note, that control experiments, where no extraction step was performed and target proteins were directly adsorbed to TEM grids, revealed that SiTF(x)=k⋆x, with constant k, for a large dynamic range (Supporting Information Figure 6.11).

6.4.4.1 Signal transfer function of apoferritin to TEM grids

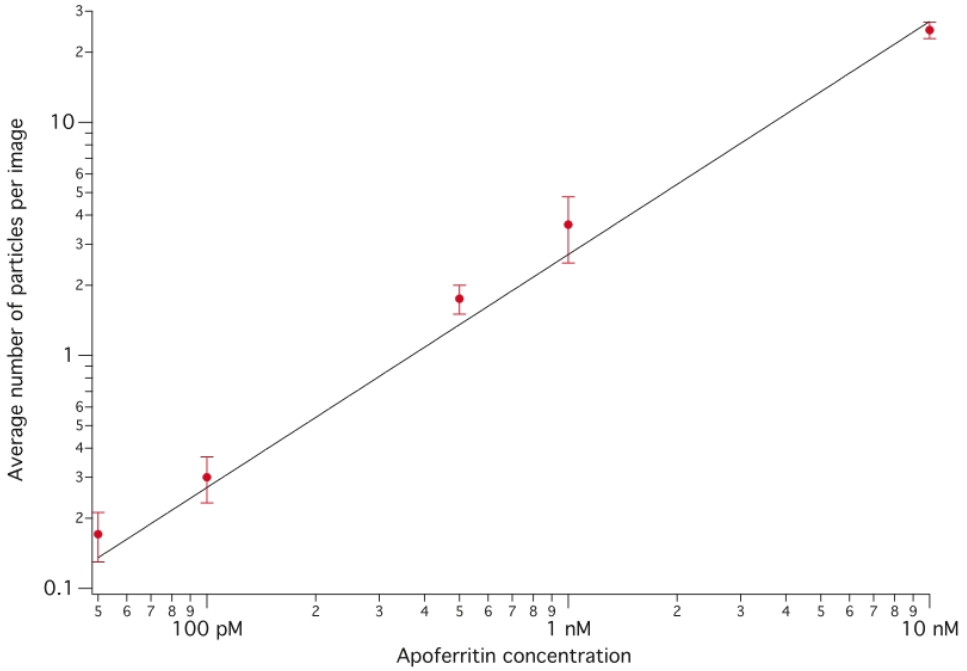


FIGURE 6.11: Signal transfer function (SiTF) of apoferritin (AF) particles to transmission electron microscopy (TEM) grids. We performed quantitative TEM (Figure 6.2) of TEM grids prepared with different concentrations of AF (input signal), ranging from 50 pM – 10 nM, to measure the amount of AF particles being adsorbed on the particular TEM grids. Here, the average number of AF particles per recorded image (PPI) for the grid preparations is shown as output signal (red dots). Error bars represent the standard deviation of the average PPIs of the five probed grid squares of the parent grid. The data documents that the SiTF function of AF particles to TEM grids is linear (black line: linear fit without offset) if no extraction step is performed ((fa(x)/x=1; see Supporting Information Figure 6.10).

6.4.5 Isolation of endogenous 20S proteasome in PBS buffer

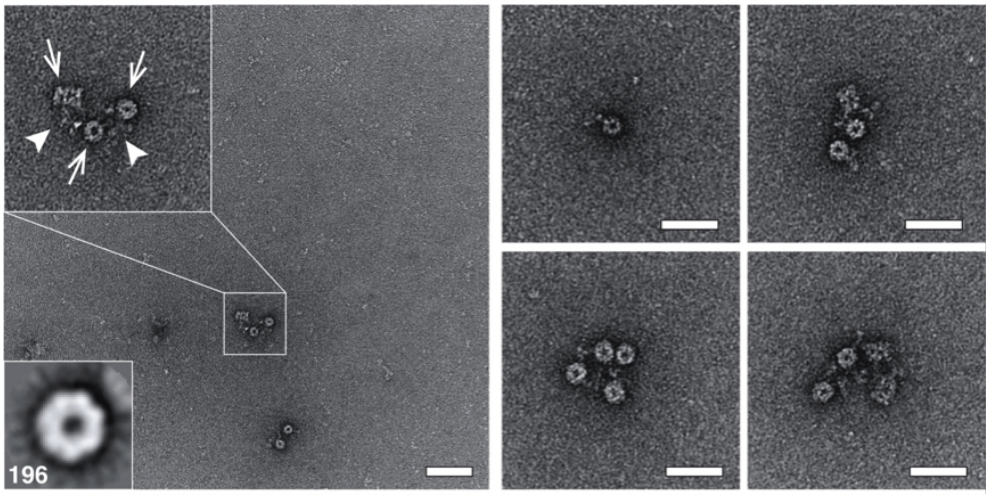


FIGURE 6.12: Isolation of endogenous 20S proteasome in PBS buffer. HEK293 cell lysate was incubated with biotinylated anti-20S proteasome antibodies (ABs). The lysate of about twenty thousand cells diluted in PBS buffer was then passed over trapped streptavidin coated magnetic beads to extract 20S proteasome, which was subsequently recovered by photocleavage and transferred onto TEM grids. Both ring-like top views and rectangular side views of the isolated 20S proteasome (arrows) are visible on the images; ABs (arrowhead) can be clearly distinguished and often link 20S proteasomes. The 2D class averages calculated (inset; bottom left, number of averaged particles) were comparable to those of purified proteasomes (Supporting Information Figure 6.13), indicating, that the presence of ABs did not interfere with 2D averaging. Note, that PBS has a higher ionic strength than HEPES buffer, which can destabilize protein-protein interactions such as the 20S-19S proteasome interaction. Indeed, we experienced that the regulatory 19S particles rarely copurify with the 20S core proteasome if PBS was used instead of HEPES (compare to Figure 6.5A&B). Scale bars, 50 nm.

6.4.6 Electron micrograph and 2D class averages of 20S proteasomes

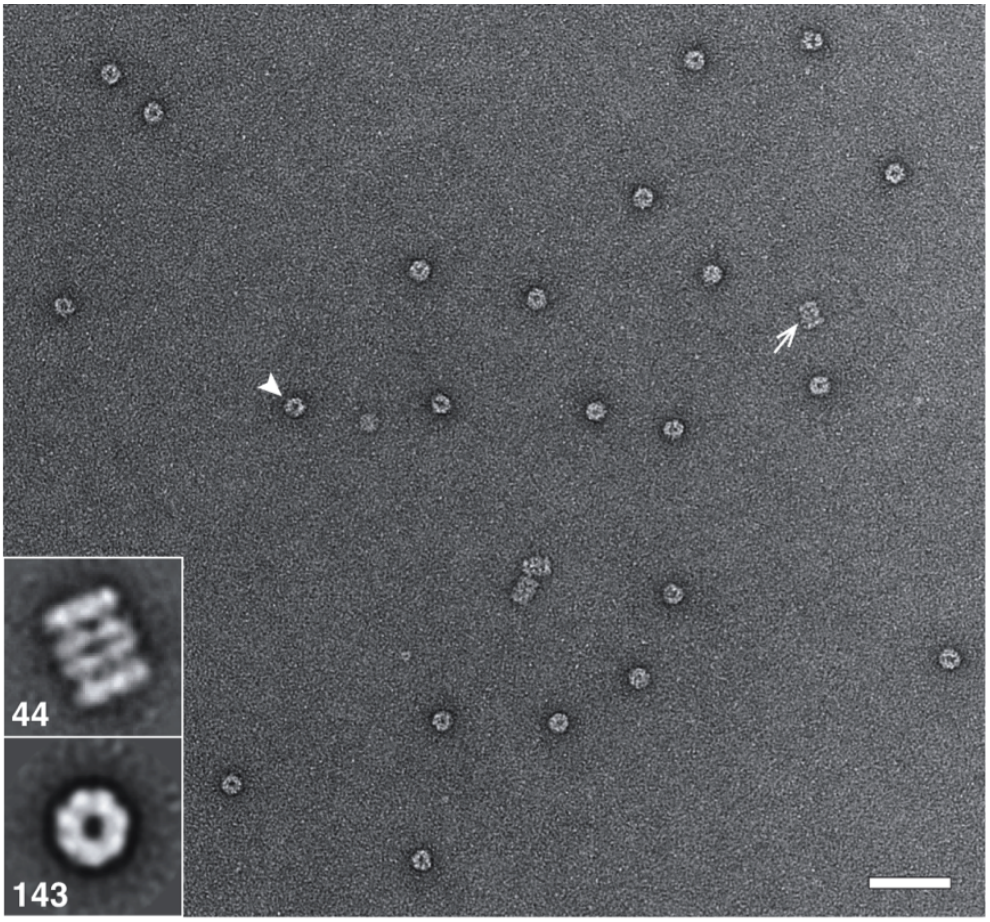


FIGURE 6.13: Negative stain transmission electron microscopy (TEM) of purified 20S proteasomes. A TEM grid was prepared by applying a 30 nM solution of purified human 20S proteasomes (Enzo Life Sciences #BML-PW8720) in PBS and imaged by TEM. Ring-shaped top views (arrowhead, 11 nm in diameter) and rectangular side views (arrow, 15 nm in length) are visible. Structural details, such as the seven-fold symmetry, can be distinguished in the calculated 2D class average (inset; bottom left, number of averaged particles). Scale bar, 50 nm.

6.4.7 COMSOL simulation of temperature rise in capillary and buffer

Figure 6.14A shows the surface temperature of the fused silica tubing after 10 min illumination. Figure 6.14B illustrates the temperature rise at the central cross-section of the capillary. The results show that the temperature rise is maximally about 0.6 K in the fused silica capillary and the liquid. However, several of the assumptions made in the model would be a worst-case scenario. Therefore, the results obtained strongly indicate that, for the experimental setup, there is no UV-light induced temperature rise in the fused silica or the PBS buffer that could affect proteins.

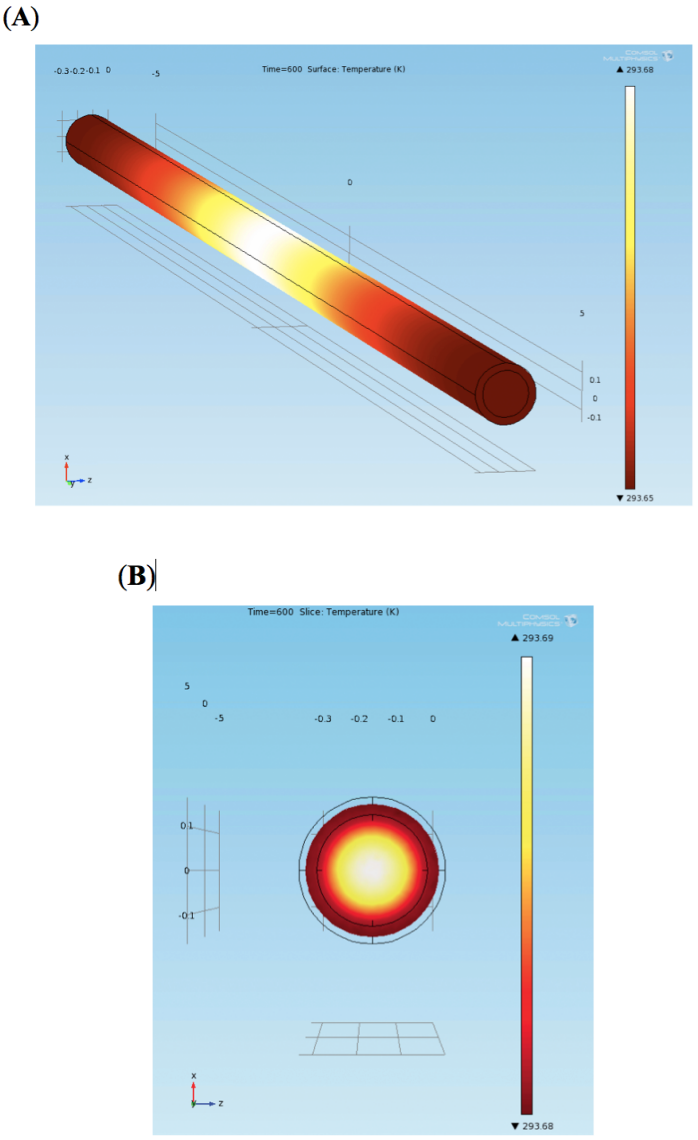


FIGURE 6.14: Finite element simulations using COMSOL Multiphysics were performed to simulate the time dependent temperature rise induced in a fused silica capillary filled with PBS buffer by an incident UV-light beam with a wavelength of 365 nm. The capillary was represented by a 10 mm long, hollow cylinder with an inner diameter of 250 μm and a wall thickness of 40 μm. The heat transfer model in COMSOL was used for modeling. The body heat load within the liquid was calculated as described in Yang et al [166]. The following assumptions were made to simplify the model: (i) the reflectivity of the fused silica is zero, (ii) there is no heat exchange with the environment and (iii) the input power of the beam does not decay due to absorption by the illuminated material. Further, the body heat load in the fused silica was neglected, as this material is almost transparent for wavelengths between 200 nm and 1 μm [167]. An absorption coefficient of 0.0016/cm, similar to that of water, was used for PBS buffer [168]. Other material properties used in the simulation were from the COMSOL material library. The incident UV-beam was set in the z-direction and represented by a 2D-Gaussian distribution in the xy-plane with an input power of 110 mW (in the experimental setup only about 25% of the emitted 440 mW UV-light is guided to the capillary). The modeled beam dimensions (σx: 0.3 mm, σy: 0.4 mm) were smaller than the apparent dimensions (elliptical shape, 4 mm long and 3 mm wide). The initial temperature was set as 293.15 K (room temperature).

Chapter 7

Microfluidics to isolate untagged proteins from cell extracts for visual analysis by electron microscopy

Dominic Giss, Simon Kemmerling, Venkata P. Dandey, Henning Stahlberg, Thomas Braun*

The following section has been published in:

Proceedings of the 17th International Conference on Miniaturized Systems for Chemistry and Life Sciences (MicroTAS), Germany, 2013 pp. 1785-1787, ISBN 978-0-9798064-6-9

1. *Center for Cellular Imaging and NanoAnalytics (C-CINA), Biozentrum,*
2. *Biozentrum, University of Basel, Klingelbergstrasse 70, 4056 Basel, Switzerland*

*Corresponding authors:Thomas.braun@unibas.ch
My contribution to this study was to implement the semi automate EM tool chain for data collection of negatively stained sample and its data processing for quantitative analysis at single molecule level.

Abstract

A simple microfluidic method for the isolation and analysis of endogenous levels of untagged protein complexes from minute amounts of cell lysate is presented. The method is based on antibodies (ABs) that are conjugated via a photocleavable linker to magnetic beads being trapped in microcapillaries by magnetic field gradients. Target proteins can be released together with their trapping ABs by photocleavage for their subsequent analysis, e.g., by transmission electron microscopy (TEM). The advantages of elution by photocleavage over classical methods are: (i) Mild recovery of proteins without changing the physiological environment and, (ii) nonspecifically adsorbed proteins are only minimally eluted.

7.1 Introduction

The investigation of multimolecular protein complexes in terms of their architecture, conformation, structure and temporal persistence is a key element of experimental systems biology. However, to date, the instrumentation for the analysis of these complexes is still poorly developed, in particular, the isolation and subsequent analysis of such complexes. Several obstacles must be overcome for these studies, e.g., the quaternary structure of these protein assemblies must be maintained and, due to the heterogeneous composition of these complexes, the analysis should be performed at single molecule level. Here we present a fast and simple microfluidic method for the isolation of large multimolecular protein complexes from cell lysate and the subsequent visual analysis of the assemblies at the single molecule level by TEM. The method allows the isolation under mild and physiological conditions and exhibits only minimal background contamination.

7.2 Functional principle

Streptavidin coated magnetic beads were trapped in a fused silica capillary by two external magnets Figure 7.1A&B. ABs, conjugated to photocleavable biotin crosslinkers, were mixed with cell lysate to bind the target protein and then immobilized on the trapped streptavidin-beads. Subsequent washing of the capillary separated the immobilized target structures from other contaminants Figure 7.1B. Recovery of proteins was achieved by illumination of the capillary with UV-light Figure 7.1C followed by elution of proteins for analysis, e.g., by TEM.

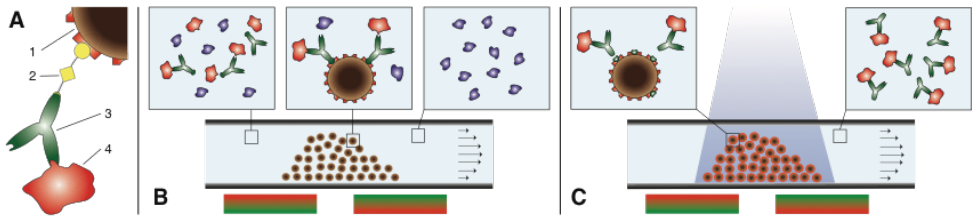


FIGURE 7.1: Working principle of the method. (A) Scheme of the composite material used for protein extraction. (1) Streptavidin coated magnetic bead; (2) photocleavable NHS-Biotin crosslinker; (3) antibody against target structure; (4) target structure. (B) Affinity extraction of protein-antibody conjugates. (C) Recovery of immobilized target proteins by photocleavage.

7.3 Experimental

ABs were biotinylated by incubating with a 10-fold molar excess of photocleavable NHS-biotin crosslinker (Ambergen, USA) for 1.5 h at pH 8.2. The reaction mixture (50 μl) was then dialyzed over night against 2 l of PBS using dialysis buttons (13 kDa cut-off) to remove unbound crosslinker. HEK293 cells were lysed using a tip sonicator and diluted to a concentration of 10000 - 35000 cells/μl. A 250 μm inner-diameter fused silica capillary (BGB Analitik AG #TSP-250350, Switzerland) was guided through a lens tube construction cube (Thorlabs #SM1C6, Germany) providing protection from UV-light and easy mounting of optical components (Figure 7-2A). About 40*106 streptavidin coated magnetic beads (Dynabeads® MyOne™, Invitrogen #656-01, Switzerland) of 1 μm diameter were loaded into the capillary at a flow rate of 20 μl/min and captured in the center of the cube by a magnetic trap consisting of two external permanent magnets (Supermagnete #Q-20-10-05-N, Switzerland). Typically, a bead plug with a length of 3 – 4 mm was created (Figure 7.2B&C). UV-light at 365 nm wavelength was emitted by a high power UV LED (Thorlabs #M365L2, Germany) mounted on top of the cube and focused onto the magnetic beads. To bind the target protein, biotinylated ABs against target structures were incubated for one hour with cell lysate. 2 μl of the AB / cell lysate mixture was flown over the trapped beads for 15 min. Thereby ABs could bind to the beads for 60 s. After the affinity extraction step, the capillary was rinsed with 200 μl of buffer at flow rates of 20 – 40 μl/min to separate unbound proteins from immobilized targets. Immobilized proteins were released by illumination of the capillary for 10 min and recovered by subsequent elution of released targets in 6 – 8 μl of buffer. The eluted protein was then transferred onto TEM sample carriers and analyzed by TEM.

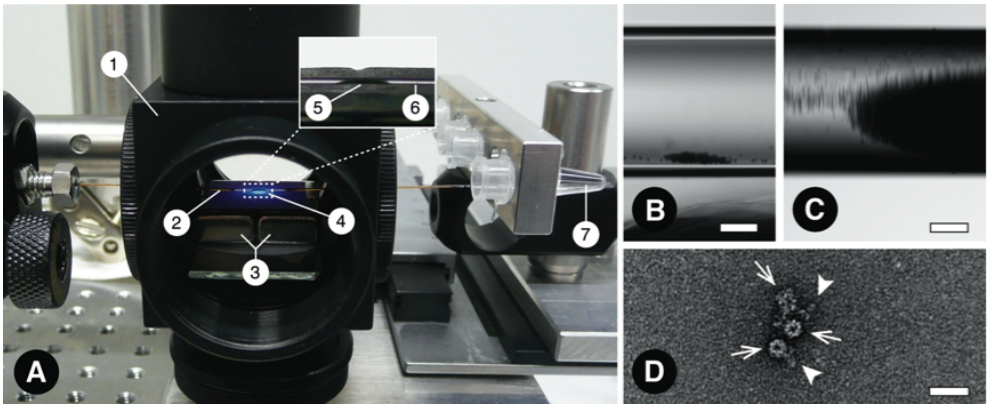


FIGURE 7.2: A: Experimental setup used for the experiments. Inset: higher magnification showing a magnetic bead plug. (1) Lens tube construction tube; (2) fused silica capillary (250 μm ID); (3) permanent magnets, building a magnetic trap; (4) UV-light of 365 nm wavelength, focused onto the magnetic bead plug; (5) magnetic bead plug of 3 – 4 mm length; (6) capillary without outer polymer coating; (7) sample container mounted on x,y-stage. B and C: The formation of a bead plug by a magnetic trap, built by the magnetic field gradient of a permanent magnet. Scale bars, 50 μm. D: Negative stain electron micrograph of endogenous 20S proteasome (arrow; anti-20S proteasome antibody, triangle) that was isolated from HEK293 cell lysate. Scale bar, 25 nm.

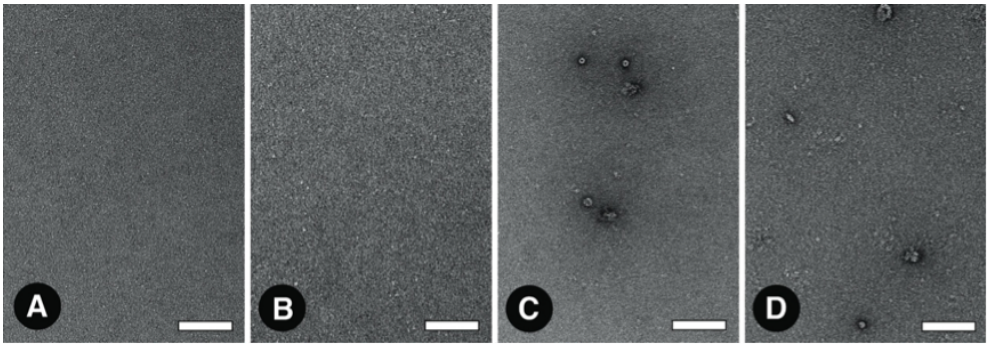


FIGURE 7.3: Comparison of contaminant background of photocleavage and competitive elution by biotin and imidazole. 3 μl of the lysate of approximately 105 cells were incubated for 15 min on a magnetic bead plug (Figure 7.2). Subsequently, different elution procedures were studied by TEM. (A) Negative control using PBS buffer. (B) Eluate obtained in PBS by illumination of the beads for 10 min with UV-light. (C) Eluate obtained by rinsing the beads for 5 min with 0.5 mM biotin in PBS buffer. D: Eluate obtained by rinsing the beads for 5 min with 500 mM imidazole in PBS buffer. Scale bars, 100 nm.

7.4 Results and discussion

The method enabled the isolation of endogenous levels of 20S proteasomes from twenty thousand HEK293 cells (Figure 7.2D). The concentration of contaminants in the resulting eluates was extremely low, as indicated by the clean background observed in TEM preparations. The recovery of immobilized proteins by photocleavage is more specific than classical chemical procedures, e.g., by competitive elution. Proteins that are nonspecifically bound to the beads or the capillary

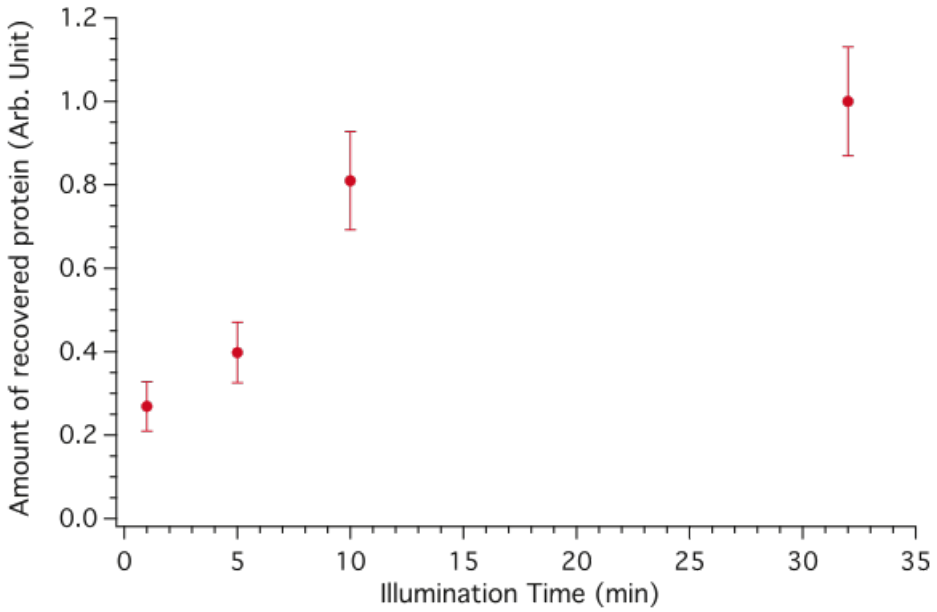


FIGURE 7.4: Effect of illumination time on the recovery of protein. Equal quantities of magnetic beads, coated with biotinylated anti-ferritin AB-apoferritin conjugates, were loaded (Figure 7.2) and illuminated for different time periods. The amount of eluted conjugates was analyzed employing quantitative TEM. The average amount of recovered particles is shown. Error bars represent the standard deviation illustrating the distribution on the particular TEM sample carrier.

surface are not significantly released during the photocleavage step (Figure 7.3). However, the change in the buffer conditions during elution with biotin or imidazole, both often used for competitive elution, significantly released contaminants from the beads and the microcapillary surface. Time series varying the illumination time for the "photoelution" showed a saturation behavior and the amount of recovered protein only slightly increased beyond 10 min of illumination (Figure 7.4). Subsequent competitive elution after 10 min of illumination, e.g. by biotin solutions, did not release more target protein. We estimate the overall isolation efficiency of the method to be around 10 – 20%. Finite element simulations estimate the UV-light induced temperature rise in the fused silica capillary and liquid to be lower than 0.6 K for 10 min of illumination, which usually does not affect protein structures. The elution by photocleavage enables mild and unperturbed recovery of proteins conserving their structural integrity since physiological conditions can be easily maintained and proteins usually do not absorb at the wavelength employed for photocleavage

7.5 Conclusion

The method presented combines magnetic beads being trapped in microcapillaries, with the usage of photocleavable crosslinkers, for unperturbed and specific recovery of target structures. Nonspecifically bound proteins are eluted only at minimal concentrations. This is crucial in microfluidics exhibiting high surface to volume ratios and the visual analysis by TEM. Using TEM, the resulting eluates can be investigated for the composition, conformation and the structure of individual protein assemblies. In future, the presented method will be combined with our new microfluidic sample deposition technique for TEM [151] and a method for the lysis of single cells. [150] The former technique enables a very efficient transfer of proteins onto TEM sample carriers, which will significantly lower the detection limit of the presented method, offering the potential even the lysates of single cells to be processed.

Chapter 8

Conclusion

This dissertation was mainly commited to image processing of PYD helical filaments and SIR complex single particle images. The techniques applied are based on the fruitful single particle reconstruction (SPR) techniques. The Chapter 2 and Chapter 3 explains about the image processing pipeline for single particle reconstruction and real space reconstruction of helical images respectively.

To complement the results of solid state NMR and solution NMR technique, cryo-EM and helical image processing are chosen for obtaining high resolution 3D model of PYD filaments. The recent structural model of the human ASC-PYD filament obtained by single particle cryo-EM showed how modern cryo-EM approaches are capable to visualize large insoluble assemblies with unprecedented level of details [116]. After tedious sample optimization, high quality EM images of the filamets are obtained with in two days of data collection using state-of-the-art electron microscope (TITAN) and direct electron detectors (K2 summit) followed by on the fly drift correction using 2dx-automator [125] software which uses the alogorithm implemented in [126]. These images diffracted to 3Å and avoided intrinsic ambiguities in determining the initial guess symmetry parameters by indexing which is crucial for Iterative Helical Real Space Reconstruction (IHRSR) technique [27]. These symmetry parameters are applied during the refinement process of IHRSR algorithm to obtain the PYD filament structure to $\sim$ 3.Å. The images of Direct Electron Detector improved the quality of the data to reach high resolution with cryo-EM. In future this advanced hardware and software in cryo-EM may help to reconstruct many flexible polymers at near atomic resolution, including the polymers whose helical symmetry cannot be predicted with certainty. Still there is a chance of improvements to obtain near atomic resolution with cryo-EM. Our final reconstruction is from $\sim$20000 overlapping

segments,each shifted 16 pixels. This corresponds to a single filament of 21.4μ length at 1.34Å/px. Comparing to single particle reconstruction we used 45000 molecules from 15000 rings of three sub units,where the spacing of the rings is 14Å rise per trimer. For a typical single particle reconstruction project there is need of 150000 particles at minimum 1Å/px to reach $\sim$ 3Å and then limiting the optical resolution may be significant due to beam tilt, curvature of Ewald sphere which results in defocus gradient across large particles and also,inaccurate estimation of CTF and alignment during image processing.

The model of the human ASC-PYD fibril at a resolution of 3.8Å describes the overall structural features of the filament and allows to visualize densities defining the pitch of α-helices and few bulky side chains on single ASC-PYD sub-units. However, additional approach to observe individual nuclei in molecular structures and their integration are highly indicated, in particular for situations where the global or local resolution of cryo-EM may not allow resolving all structural details of interest. Furthermore, the combination with dynamics measurements allows a complementary view on the protein structure and function. Recently, the use of a molecular modeling protocol to combine solid state NMR restraints and cryo-EM data proved the potential of such hybrid approach to render atomic model of large macro-molecular assembly [112]. Our collaborators were able to integrate complementary structural information from cryo-EM and solid state NMR into an ab initio structure calculation procedure to gain the atomic resolution structure of the mouse ASC-PYD filament. The availability of the mouse structure will provide important for in vivo studies.

For SIR complex structural studies, several interactions between the SIR complex and binding partners have been mapped and structural information of almost 50% of the complex [86, 87, 95]are available prior to this thesis. However, even though we have a thorough understanding of the genetic requirements for silencing, we do not know how the SIR complex binds chromatin and silences transcription. Our main aim for this study is obtain a medium resolution ($\sim$ 16Å) model to fit the X-ray structure of the complex. Due to oligomerization properties and flexibility of the SIR complex limited us to obtain a reliable 3D model. The isolation of the smallest sub units (Sir2-Sir4 hetero tetramer,Sir3 dimer and Sir2-Sir4-AB mutant forming dimers) and semi-automated EM tool chain only helped to understand size variations with in the complexes. Our investigation using the RCT reconstruction technique revealed the heterogeneity of the sample. The resulting class averages had various shapes, so that we could not obtain a clear 3D volume. A major bottleneck for EM analysis is to obtain homogeneous particles, which need to

be optimized for this complex. Structural information employing negative stain TEM is very limited to about 2nm and so, either MBP- cryo-EM analysis of the homogeneous particles at high concentration using state-of-the-art direct electron detector can help to reach near atomic resolution.

Bibliography

[1] E. V. Orlova and H. R. Saibil. Methods for three-dimensional reconstruction of heterogeneous assemblies. *Meth. Enzymol.*, 482:321–341, 2010.

[2] M. Halic, T. Becker, M. R. Pool, C. M. Spahn, R. A. Grassucci, J. Frank, and R. Beckmann. Structure of the signal recognition particle interacting with the elongation-arrested ribosome. *Nature*, 427(6977):808–814, Feb 2004.

[3] J. B. Heymann, N. Cheng, W. W. Newcomb, B. L. Trus, J. C. Brown, and A. C. Steven. Dynamics of herpes simplex virus capsid maturation visualized by time-lapse cryo-electron microscopy. *Nat. Struct. Biol.*, 10(5):334–341, May 2003.

[4] M. Valle, A. Zavialov, J. Sengupta, U. Rawat, M. Ehrenberg, and J. Frank. Locking and unlocking of ribosomal motions. *Cell*, 114(1):123–134, Jul 2003.

[5] B. P. Klaholz, A. G. Myasnikov, and M. Van Heel. Visualization of release factor 3 on the ribosome during termination of protein synthesis. *Nature*, 427(6977):862–865, Feb 2004.

[6] N. Fischer, A. L. Konevega, W. Wintermeyer, M. V. Rodnina, and H. Stark. Ribosome dynamics and tRNA movement by time-resolved electron cryomicroscopy. *Nature*, 466(7304):329–333, Jul 2010.

[7] E. V. Orlova and H. R. Saibil. Structural analysis of macromolecular assemblies by electron microscopy. *Chem. Rev.*, 111(12):7710–7748, Dec 2011.

[8] N. Elad, D. K. Clare, H. R. Saibil, and E. V. Orlova. Detection and separation of heterogeneity in molecular complexes by statistical analysis of their two-dimensional projections. *J. Struct. Biol.*, 162(1):108–120, Apr 2008.

[9] N. Elad, G. W. Farr, D. K. Clare, E. V. Orlova, A. L. Horwich, and H. R. Saibil. Topologies of a substrate protein bound to the chaperonin GroEL. *Mol. Cell*, 26(3):415–426, May 2007.

[10] H. E. White, E. V. Orlova, S. Chen, L. Wang, A. Ignatiou, B. Gowen, T. Stromer, T. M. Franzmann, M. Haslbeck, J. Buchner, and H. R. Saibil. Multiple distinct assemblies reveal conformational flexibility in the small heat shock protein Hsp26. *Structure*, 14(7):1197–1204, Jul 2006.

[11] H. E. White, H. R. Saibil, A. Ignatiou, and E. V. Orlova. Recognition and separation of single particles with size variation by statistical analysis of their images. *J. Mol. Biol.*, 336(2):453–460, Feb 2004.

[12] P. A. Penczek, C. Yang, J. Frank, and C. M. Spahn. Estimation of variance in single-particle reconstruction using the bootstrap technique. *J. Struct. Biol.*, 154(2):168–183, May 2006.

[13] T. Wagenknecht, J. M. Carazo, M. Radermacher, and J. Frank. Three-dimensional reconstruction of the ribosome from Escherichia coli. *Biophys. J.*, 55(3):455–464, Mar 1989.

[14] M. Radermacher. Three-dimensional reconstruction of single particles from random and nonrandom tilt series. *J Electron Microsc Tech*, 9(4):359–394, Aug 1988.

[15] M. Radermacher, V. Rao, R. Grassucci, J. Frank, A. P. Timerman, S. Fleischer, and T. Wagenknecht. Cryo-electron microscopy and three-dimensional reconstruction of the calcium release channel/ryanodine receptor from skeletal muscle. *J. Cell Biol.*, 127(2):411–423, Oct 1994.

[16] J. J. Fernandez, S. Li, and R. A. Crowther. CTF determination and correction in electron cryotomography. *Ultramicroscopy*, 106(7):587–596, May 2006.

[17] A. E. Leschziner and E. Nogales. The orthogonal tilt reconstruction method: an approach to generating single-class volumes with no missing cone for ab initio reconstruction of asymmetric particles. *J. Struct. Biol.*, 153(3):284–299, Mar 2006.

[18] M. Radermacher, T. Wagenknecht, A. Verschoor, and J. Frank. Three-dimensional reconstruction from a single-exposure, random conical tilt series applied to the 50S ribosomal subunit of Escherichia coli. *J Microsc*, 146(Pt 2):113–136, May 1987.

[19] M. Van Heel. Angular reconstitution: a posteriori assignment of projection directions for 3D reconstruction. *Ultramicroscopy*, 21(2):111–123, 1987.

[20] E. H. Egelman. Reducing irreducible complexity: divergence of quaternary structure and function in macromolecular assemblies. *Curr. Opin. Cell Biol.*, 22(1):68–74, Feb 2010.

[21] P. B. Moore, H. E. Huxley, and D. J. DeRosier. Three-dimensional reconstruction of F-actin, thin filaments and decorated thin filaments. *J. Mol. Biol.*, 50(2):279–295, Jun 1970.

[22] D. J. De Rosier and A. Klug. Reconstruction of three dimensional structures from electron micrographs. *Nature*, 217(5124):130–134, Jan 1968.

[23] A. Klug, F. H. C. Crick, and H. W. Wyckoff. Diffraction by helical structures. *Acta Crystallographica*, 11(3):199–213, Mar 1958. doi: 10.1107/S0365110X58000517. URL http://dx.doi.org/10.1107/S0365110X58000517.

[24] M. Stewart. Computer image processing of electron micrographs of biological structures with helical symmetry. *J Electron Microsc Tech*, 9(4):325–358, Aug 1988.

[25] E. H. Egelman. A robust algorithm for the reconstruction of helical filaments using single-particle methods. *Ultramicroscopy*, 85(4):225–234, Dec 2000.

[26] E. H. Egelman. Single-particle reconstruction from EM images of helical filaments. *Curr. Opin. Struct. Biol.*, 17(5):556–561, Oct 2007.

[27] E. H. Egelman. The iterative helical real space reconstruction method: surmounting the problems posed by real polymers. *J. Struct. Biol.*, 157(1): 83–94, Jan 2007.

[28] H. A. Al-Khayat, E. P. Morris, and J. M. Squire. The 7-stranded structure of relaxed scallop muscle myosin filaments: support for a common head configuration in myosin-regulated muscles. *J. Struct. Biol.*, 166(2):183–194, May 2009.

[29] E. H. Egelman. An algorithm for straightening images of curved filamentous structures. *Ultramicroscopy*, 19(4):367–373, 1986.

[30] T. Fujii, T. Kato, and K. Namba. Specific arrangement of alpha-helical coiled coils in the core domain of the bacterial flagellar hook for the universal joint function. *Structure*, 17(11):1485–1493, Nov 2009.

[31] E. H. Egelman, N. Francis, and D. J. DeRosier. F-actin is a helix with a random variable twist. *Nature*, 298(5870):131–135, Jul 1982.

[32] C. Suloway, J. Pulokas, D. Fellmann, A. Cheng, F. Guerra, J. Quispe, S. Stagg, C. S. Potter, and B. Carragher. Automated molecular microscopy: the new Leginon system. *J. Struct. Biol.*, 151(1):41–60, Jul 2005.

[33] G. Tang, L. Peng, P. R. Baldwin, D. S. Mann, W. Jiang, I. Rees, and S. J. Ludtke. EMAN2: an extensible image processing suite for electron microscopy. *J. Struct. Biol.*, 157(1):38–46, Jan 2007.

[34] B. K. Rath and J. Frank. Fast automatic particle picking from cryo-electron micrographs using a locally normalized cross-correlation function: a case study. *J. Struct. Biol.*, 145(1-2):84–90, 2004.

[35] M. van Heel and J. Frank. Use of multivariate statistics in analysing the images of biological macromolecules. *Ultramicroscopy*, 6(2):187–194, 1981.

[36] M. van Heel. Multivariate statistical classification of noisy images (randomly oriented biological macromolecules). *Ultramicroscopy*, 13(1-2):165–183, 1984.

[37] P. A. Penczek, J. Zhu, and J. Frank. A common-lines based method for determining orientations for N ¿ 3 particle projections simultaneously. *Ultramicroscopy*, 63(3-4):205–218, 1996.

[38] M. Hohn, G. Tang, G. Goodyear, P. R. Baldwin, Z. Huang, P. A. Penczek, C. Yang, R. M. Glaeser, P. D. Adams, and S. J. Ludtke. SPARX, a new environment for Cryo-EM image processing. *J. Struct. Biol.*, 157(1):47–55, Jan 2007.

[39] M. van Heel and G. Harauz. Biological macromolecules explored by pattern recognition. *Scanning Microsc. Suppl.*, 2:295–301, 1988.

[40] M. van Heel, G. Harauz, E. V. Orlova, R. Schmidt, and M. Schatz. A new generation of the IMAGIC image processing system. *J. Struct. Biol.*, 116 (1):17–24, 1996.

[41] F. Forster, S. Pruggnaller, A. Seybert, and A. S. Frangakis. Classification of cryo-electron sub-tomograms using constrained correlation. *J. Struct. Biol.*, 161(3):276–286, Mar 2008.

[42] W. Hoppe, R. Langer, G. Knesch, and C. Poppe. [Protein crystal structure analysis with electron radiation]. *Naturwissenschaften*, 55(7):333–336, Jul 1968.

[43] V. Knauer, R. Hegerl, and W. Hoppe. Three-dimensional reconstruction and averaging of 30 S ribosomal subunits of Escherichia coli from electron micrographs. *J. Mol. Biol.*, 163(3):409–430, Jan 1983.

[44] J. Walz, D. Typke, M. Nitsch, A. J. Koster, R. Hegerl, and W. Baumeister. Electron Tomography of Single Ice-Embedded Macromolecules: Three-Dimensional Alignment and Classification. *J. Struct. Biol.*, 120(3):387–395, Dec 1997.

[45] Lebart. Correspondence analysis and classification. *Seventh International Conference on Multivariate Analysis*, page 347, september 1992.

[46] T. Kohonen. Cortical maps. *Nature*, 346(6279):24, Jul 1990.

[47] R. Marabini and J. M. Carazo. Pattern recognition and classification of images of biological macromolecules using artificial neural networks. *Biophys. J.*, 66(6):1804–1814, Jun 1994.

[48] Z. Zhao. Class Averaging in Cryo-EM Single Particle Reconstruction, 2015. URL http://arks.princeton.edu/ark:/88435/dsp01wp988j94j. Accessed:2013.

[49] Frank Natterer. *The Mathematics of Computerized Tomography*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2001. ISBN 0-89871-493-1.

[50] P. A. Penczek, J. Zhu, and J. Frank. A common-lines based method for determining orientations for N 3 particle projections simultaneously. *Ultramicroscopy*, 63(3-4):205–218, Jul 1996.

[51] M. van Heel, B. Gowen, R. Matadeen, E. V. Orlova, R. Finn, T. Pape, D. Cohen, H. Stark, R. Schmidt, M. Schatz, and A. Patwardhan. Single-particle electron cryo-microscopy: towards atomic resolution. *Q. Rev. Biophys.*, 33 (4):307–369, Nov 2000.

[52] J. Frank. Averaging of low exposure electron micrographs of non-periodic objects. *Ultramicroscopy*, 1(2):159–162, Dec 1975.

[53] W. O. Saxton and J. Frank. Motif detection in quantum noise-limited electron micrographs by cross-correlation. *Ultramicroscopy*, 2(2-3):219–227, Apr 1977.

[54] E. J. Boekema, J. A. Berden, and M. G. van Heel. Structure of mitochondrial F1-ATPase studied by electron microscopy and image processing. *Biochim. Biophys. Acta*, 851(3):353–360, Oct 1986.

[55] M. Van Heel, J. P. Bretaudiere, and Frank J. Classification and multireference alignment of images of macromolecules. *In Proceedings of the 10th International Congress on Electron Microscopy*, 1:563–564, 1982.

[56] N. Grigorieff. Resolution measurement in structures derived from single particles. *Acta Crystallogr. D Biol. Crystallogr.*, 56(Pt 10):1270–1277, Oct 2000.

[57] P. Penczek, M. Radermacher, and J. Frank. Three-dimensional reconstruction of single particles embedded in ice. *Ultramicroscopy*, 40(1):33–53, Jan 1992.

[58] J. Frank, M. Radermacher, P. Penczek, J. Zhu, Y. Li, M. Ladjadj, and A. Leith. SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields. *J. Struct. Biol.*, 116(1):190–199, 1996.

[59] Lisa Borland and Marin van Heel. Classification of image data in conjugate representation spaces. *J. Opt. Soc. Am. A*, 7(4):601–610, Apr 1990. doi: 10.1364/JOSAA.7.000601. URL http://josaa.osa.org/abstract.cfm?URI=josaa-7-4-601.

[60] M. van Heel, M. Schatz, and E Orlova. Correlation functions revisited. *Ultramicroscopy*, 46(1-4):307–316, 1992.

[61] M. Radermacher, T. Wagenknecht, A. Verschoor, and J. Frank. Three-dimensional structure of the large ribosomal subunit from Escherichia coli. *EMBO J.*, 6(4):1107–1114, Apr 1987.

[62] C. Yoshioka, J. Pulokas, D. Fellmann, C. S. Potter, R. A. Milligan, and B. Carragher. Automation of random conical tilt and orthogonal tilt data collection using feature-based correlation. *J. Struct. Biol.*, 159(3):335–346, Sep 2007.

[63] X. Liu and H. W. Wang. Single particle electron microscopy reconstruction of the exosome complex using the random conical tilt method. *J Vis Exp*, (49), 2011.

[64] Zhao, Zhizhen, Singer, and Amit. Rotationally invariant image representation for viewing direction classification in cryo-em. *Journal of Structural Biology*, (1):153–166. doi: 10.1016/j.jsb.2014.03.003.

[65] M. Radermacher. Three-dimensional reconstruction from random projections: orientational alignment via Radon transforms. *Ultramicroscopy*, 53 (2):121–136, Feb 1994.

[66] Michael Radermacher and Teresa Ruiz. Three-dimensional reconstruction of single particles in electron microscopy. 319, November 2005. URL http://www.springerprotocols.com/Abstract/doi/10.1007/978-1-59259-993-6_20.

[67] S. Lanzavecchia, P. L. Bellon, and M. Radermacher. Fast and accurate three-dimensional reconstruction from projections with random orientations via radon transforms. *J. Struct. Biol.*, 128(2):152–164, Dec 1999.

[68] *The Radon Transform and Some of Its Applications*. Krieger Publishing Company. ISBN 0894647180.

[69] E. Behrmann, G. Tao, D. L. Stokes, E. H. Egelman, S. Raunser, and P. A. Penczek. Real-space processing of helical filaments in SPARX. *J. Struct. Biol.*, 177(2):302–313, Feb 2012.

[70] B. Wu, A. Peisley, D. Tetrault, Z. Li, E. H. Egelman, K. E. Magor, T. Walz, P. A. Penczek, and S. Hur. Molecular imprinting as a signal-activation mechanism of the viral RNA sensor RIG-I. *Mol. Cell*, 55(4):511–523, Aug 2014.

[71] R. D. Kornberg. Structure of chromatin. *Annu. Rev. Biochem.*, 46:931–954, 1977.

[72] J. D. McGhee, D. C. Rau, E. Charney, and G. Felsenfeld. Orientation of the nucleosome within the higher order structure of chromatin. *Cell*, 22(1 Pt 1):87–96, Nov 1980.

[73] K. Luger, T. J. Rechsteiner, A. J. Flaus, M. M. Waye, and T. J. Richmond. Characterization of nucleosome core particles containing histone proteins made in bacteria. *J. Mol. Biol.*, 272(3):301–311, Sep 1997.

[74] M. Grunstein. Nucleosomes: regulators of transcription. *Trends Genet.*, 6 (12):395–400, Dec 1990.

[75] S. M. Paranjape, R. T. Kamakaka, and J. T. Kadonaga. Role of chromatin structure in the regulation of transcription by RNA polymerase II. *Annu. Rev. Biochem.*, 63:265–297, 1994.

[76] L. Hong, S. Brautigam, and A. Rethwilm. Expression of the human foamy virus bel-1 transactivator in insect cells. *Virus Res.*, 30(1):89–95, Oct 1993.

[77] D. J. Steger and J. L. Workman. Remodeling chromatin structures for transcription: what happens to the histones? *Bioessays*, 18(11):875–884, Nov 1996.

[78] T. S. Woodcock, E. E. Boyle, R. E. Roughley, P. G. Kevan, R. N. Labbee, A. B. Smith, H. Goulet, D. Steinke, and S. J. Adamowicz. The diversity and biogeography of the Coleoptera of Churchill: insights from DNA barcoding. *BMC Ecol.*, 13:40, 2013.

[79] J. Rine and I. Herskowitz. Four genes responsible for a position effect on expression from HML and HMR in Saccharomyces cerevisiae. *Genetics*, 116 (1):9–22, May 1987.

[80] L. N. Rusche, A. L. Kirchmaier, and J. Rine. The establishment, inheritance, and function of silenced chromatin in Saccharomyces cerevisiae. *Annu. Rev. Biochem.*, 72:481–516, 2003.

[81] S. Strahl-Bolsinger, A. Hecht, K. Luo, and M. Grunstein. SIR2 and SIR4 interactions differ in core and extended telomeric heterochromatin in yeast. *Genes Dev.*, 11(1):83–93, Jan 1997.

[82] A. Hecht, S. Strahl-Bolsinger, and M. Grunstein. Spreading of transcriptional repressor SIR3 from telomeric heterochromatin. *Nature*, 383(6595): 92–96, Sep 1996.

[83] L. Maillet, C. Boscheron, M. Gotta, S. Marcand, E. Gilson, and S. M. Gasser. Evidence for silencing compartments within the yeast nucleus: a role for telomere proximity and Sir protein concentration in silencer-mediated repression. *Genes Dev.*, 10(14):1796–1811, Jul 1996.

[84] S. Marcand, S. M. Gasser, and E. Gilson. Chromatin: a sticky silence. *Curr. Biol.*, 6(10):1222–1225, Oct 1996.

[85] F. van Leeuwen, P. R. Gafken, and D. E. Gottschling. Dot1p modulates silencing in yeast by methylation of the nucleosome core. *Cell*, 109(6):745–756, Jun 2002.

[86] F. Martino, S. Kueng, P. Robinson, M. Tsai-Pflugfelder, F. van Leeuwen, M. Ziegler, F. Cubizolles, M. M. Cockell, D. Rhodes, and S. M. Gasser. Reconstitution of yeast silent chromatin: multiple contact sites and O-AADPR binding load SIR complexes onto nucleosomes in vitro. *Mol. Cell*, 33(3):323–334, Feb 2009.

[87] M. Oppikofer, S. Kueng, F. Martino, S. Soeroes, S. M. Hancock, J. W. Chin, W. Fischle, and S. M. Gasser. A dual role of H4K16 acetylation in the establishment of yeast silent chromatin. *EMBO J.*, 30(13):2610–2621, Jul 2011.

[88] M. Onishi, G. G. Liou, J. R. Buchberger, T. Walz, and D. Moazed. Role of the conserved Sir3-BAH domain in nucleosome binding and silent chromatin assembly. *Mol. Cell*, 28(6):1015–1028, Dec 2007.

[89] J. R. Buchberger, M. Onishi, G. Li, J. Seebacher, A. D. Rudner, S. P. Gygi, and D. Moazed. Sir3-nucleosome interactions in spreading of silent chromatin in Saccharomyces cerevisiae. *Mol. Cell. Biol.*, 28(22):6903–6918, Nov 2008.

[90] V. Sampath, P. Yuan, I. X. Wang, E. Prugar, F. van Leeuwen, and R. Sternglanz. Mutational analysis of the Sir3 BAH domain reveals multiple points of interaction with nucleosomes. *Mol. Cell. Biol.*, 29(10):2532–2545, May 2009.

[91] K. J. Armache, J. D. Garlick, D. Canzio, G. J. Narlikar, and R. E. Kingston. Structural basis of silencing: Sir3 BAH domain in complex with a nucleosome at 3.0 Å resolution. *Science*, 334(6058):977–982, Nov 2011.

[92] S. P. Bell, J. Mitchell, J. Leber, R. Kobayashi, and B. Stillman. The multidomain structure of Orc1p reveals similarity to regulators of DNA replication and transcriptional silencing. *Cell*, 83(4):563–568, Nov 1995.

[93] A. F. Neuwald, L. Aravind, J. L. Spouge, and E. V. Koonin. AAA+: A class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes. *Genome Res.*, 9(1):27–43, Jan 1999.

[94] J. J. Connelly, P. Yuan, H. C. Hsu, Z. Li, R. M. Xu, and R. Sternglanz. Structure and function of the Saccharomyces cerevisiae Sir3 BAH domain. *Mol. Cell. Biol.*, 26(8):3256–3265, Apr 2006.

[95] S. Ehrentraut, M. Hassler, M. Oppikofer, S. Kueng, J. M. Weber, J. W. Mueller, S. M. Gasser, A. G. Ladurner, and A. E. Ehrenhofer-Murray. Structural basis for the role of the Sir3 AAA+ domain in silencing: interaction with Sir4 and unmethylated histone H3K79. *Genes Dev.*, 25(17):1835–1846, Sep 2011.

[96] H. Liaw and A. J. Lustig. Sir3 C-terminal domain involvement in the initiation and spreading of heterochromatin. *Mol. Cell. Biol.*, 26(20):7616–7631, Oct 2006.

[97] M. Oppikofer, S. Kueng, and S. M. Gasser. SIR-nucleosome interactions: structure-function relationships in yeast silent chromatin. *Gene*, 527(1): 10–25, Sep 2013.

[98] S. Kueng, M. Oppikofer, and S. M. Gasser. SIR proteins and the assembly of silent chromatin in budding yeast. *Annu. Rev. Genet.*, 47:275–306, 2013.

[99] B. Kastner, N. Fischer, M. M. Golas, B. Sander, P. Dube, D. Boehringer, K. Hartmuth, J. Deckert, F. Hauer, E. Wolf, H. Uchtenhagen, H. Urlaub, F. Herzog, J. M. Peters, D. Poerschke, R. Luhrmann, and H. Stark. GraFix: sample preparation for single-particle electron cryomicroscopy. *Nat. Methods*, 5(1):53–55, Jan 2008.

[100] A. M. Roseman. FindEM–a fast, efficient program for automatic selection of particles from electron micrographs. *J. Struct. Biol.*, 145(1-2):91–99, 2004.

[101] G. G. Liou, J. C. Tanny, R. G. Kruger, T. Walz, and D. Moazed. Assembly of the SIR complex and its regulation by O-acetyl-ADP-ribose, a product of NAD-dependent histone deacetylation. *Cell*, 121(4):515–527, May 2005.

[102] T. Ruiz and M. Radermacher. Three-dimensional analysis of single particles by electron microscopy: sample preparation and data acquisition. *Methods Mol. Biol.*, 319:403–425, 2006.

[103] K. Schroder and J. Tschopp. The inflammasomes. *Cell*, 140(6):821–832, Mar 2010.

[104] Z. Hu, C. Yan, P. Liu, Z. Huang, R. Ma, C. Zhang, R. Wang, Y. Zhang, F. Martinon, D. Miao, H. Deng, J. Wang, J. Chang, and J. Chai. Crystal structure of NLRC4 reveals its autoinhibition mechanism. *Science*, 341 (6142):172–175, Jul 2013.

[105] J. Masumoto, S. Taniguchi, K. Ayukawa, H. Sarvotham, T. Kishino, N. Niikawa, E. Hidaka, T. Katsuyama, T. Higuchi, and J. Sagara. ASC, a novel 22-kDa protein, aggregates during apoptosis of human promyelocytic leukemia HL-60 cells. *J. Biol. Chem.*, 274(48):33835–33838, Nov 1999.

[106] E. de Alba. Structure and interdomain dynamics of apoptosis-associated speck-like protein containing a CARD (ASC). *J. Biol. Chem.*, 284(47): 32932–32941, Nov 2009.

[107] E. Liepinsh, R. Barbals, E. Dahl, A. Sharipo, E. Staub, and G. Otting. The death-domain fold of the ASC PYRIN domain, presenting a basis for PYRIN/PYRIN recognition. *J. Mol. Biol.*, 332(5):1155–1163, Oct 2003.

[108] P. R. Vajjhala, R. E. Mirams, and J. M. Hill. Multiple binding sites on the pyrin domain of ASC protein allow self-association and interaction with NLRP3 protein. *J. Biol. Chem.*, 287(50):41732–41743, Dec 2012.

[109] T. Srimathi, S. L. Robbins, R. L. Dubas, H. Chang, H. Cheng, H. Roder, and Y. C. Park. Mapping of POP1-binding site on pyrin domain of ASC. *J. Biol. Chem.*, 283(22):15390–15398, May 2008.

[110] P. R. Vajjhala, S. Kaiser, S. J. Smith, Q. R. Ong, S. L. Soh, K. J. Stacey, and J. M. Hill. Identification of multifaceted binding modes for pyrin and ASC pyrin domains gives insights into pyrin inflammasome assembly. *J. Biol. Chem.*, 289(34):23504–23519, Aug 2014.

[111] C. Wasmer, A. Lange, H. Van Melckebeke, A. B. Siemer, R. Riek, and B. H. Meier. Amyloid fibrils of the HET-s(218-289) prion form a beta solenoid with a triangular hydrophobic core. *Science*, 319(5869):1523–1526, Mar 2008.

[112] A. Loquet, N. G. Sgourakis, R. Gupta, K. Giller, D. Riedel, C. Goosmann, C. Griesinger, M. Kolbe, D. Baker, S. Becker, and A. Lange. Atomic model of the type III secretion system needle. *Nature*, 486(7402):276–279, Jun 2012.

[113] M. Vilar, H. T. Chou, T. Luhrs, S. K. Maji, D. Riek-Loher, R. Verel, G. Manning, H. Stahlberg, and R. Riek. The fold of alpha-synuclein fibrils. *Proc. Natl. Acad. Sci. U.S.A.*, 105(25):8637–8642, Jun 2008.

[114] J. Greenwald, C. Buhtz, C. Ritter, W. Kwiatkowski, S. Choe, M. L. Maddelein, F. Ness, S. Cescau, A. Soragni, D. Leitz, S. J. Saupe, and R. Riek. The mechanism of prion inhibition by HET-S. *Mol. Cell*, 38(6):889–899, Jun 2010.

[115] J. X. Lu, W. Qiang, W. M. Yau, C. D. Schwieters, S. C. Meredith, and R. Tycko. Molecular structure of Î2-amyloid fibrils in Alzheimer's disease brain tissue. *Cell*, 154(6):1257–1268, Sep 2013.

[116] A. Lu, V. G. Magupalli, J. Ruan, Q. Yin, M. K. Atianand, M. R. Vos, G. F. Schroder, K. A. Fitzgerald, H. Wu, and E. H. Egelman. Unified polymerization mechanism for the assembly of ASC-dependent inflammasomes. *Cell*, 156(6):1193–1206, Mar 2014.

[117] T. Herrmann, P. Guntert, and K. Wuthrich. Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. *J. Mol. Biol.*, 319(1): 209–227, May 2002.

[118] C. D. Schwieters, J. J. Kuszewski, N. Tjandra, and G. M. Clore. The Xplor-NIH NMR molecular structure determination package. *J. Magn. Reson.*, 160(1):65–73, Jan 2003.

[119] P. Broz, J. von Moltke, J. W. Jones, R. E. Vance, and D. M. Monack. Differential requirement for Caspase-1 autoproteolysis in pathogen-induced cell death and cytokine processing. *Cell Host Microbe*, 8(6):471–483, Dec 2010.

[120] J. P. Demers, B. Habenstein, A. Loquet, S. Kumar Vasa, K. Giller, S. Becker, D. Baker, A. Lange, and N. G. Sgourakis. High-resolution structure of the Shigella type-III secretion needle by solid-state NMR and cryo-electron microscopy. *Nat Commun*, 5:4976, 2014.

[121] E. de Alba. 1H, 15N and 13C backbone and side chain chemical shifts of human ASC (apoptosis-associated speck-like protein containing a CARD domain). *Biomol NMR Assign*, 1(1):135–137, Jul 2007.

[122] T. Szyperski, S. Scheek, J. Johansson, G. Assmann, U. Seedorf, and K. Wüthrich. {NMR} determination of the secondary structure and the three-dimensional polypeptide backbone fold of the human sterol carrier protein 2. {*FEBS*} *Letters*, 335(1):18 – 26, 1993. ISSN 0014-5793. doi: http://dx.doi.org/10.1016/0014-5793(93)80431-S. URL http://www.sciencedirect.com/science/article/pii/001457939380431S.

[123] R.L.J. Keller. *Computer aided resonance assignment tutorial*. Cantina, 2004. URL http://cara.nmr.ch.

[124] X. Li, S. Q. Zheng, K. Egami, D. A. Agard, and Y. Cheng. Influence of electron dose rate on electron counting images recorded with the K2 camera. *J. Struct. Biol.*, 184(2):251–260, Nov 2013.

[125] S. Scherer, J. Kowal, M. Chami, V. Dandey, M. Arheit, P. Ringler, and H. Stahlberg. 2dx_automator-implementation of a semiautomatic high-throughput high-resolution cryo-electron crystallography pipeline. *J. Struct. Biol.*, 186(2):302–307, May 2014.

[126] X. Li, P. Mooney, S. Zheng, C. R. Booth, M. B. Braunfeld, S. Gubbens, D. A. Agard, and Y. Cheng. Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nat. Methods*, 10(6):584–590, Jun 2013.

[127] J. A. Mindell and N. Grigorieff. Accurate determination of local defocus and specimen tilt in electron microscopy. *J. Struct. Biol.*, 142(3):334–347, Jun 2003.

[128] P. Emsley and K. Cowtan. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.*, 60(Pt 12 Pt 1):2126–2132, Dec 2004.

[129] P. Guntert. Automated NMR structure calculation with CYANA. *Methods Mol. Biol.*, 278:353–378, 2004.

[130] A. Aderem. Systems biology: its practice and challenges. *Cell*, 121(4): 511–513, May 2005.

[131] S. Fields and O. Song. A novel genetic system to detect protein-protein interactions. *Nature*, 340(6230):245–246, Jul 1989.

[132] T. C. Walther and M. Mann. Mass spectrometry-based proteomics in cell biology. *J. Cell Biol.*, 190(4):491–500, Aug 2010.

[133] I. Uzoma and H. Zhu. Interactome mapping: using protein microarray technology to reconstruct diverse protein networks. *Genomics Proteomics Bioinformatics*, 11(1):18–28, Feb 2013.

[134] M. Vidal, M. E. Cusick, and A. L. Barabasi. Interactome networks and human disease. *Cell*, 144(6):986–998, Mar 2011.

[135] C. Klockenbusch and J. Kast. Optimization of formaldehyde cross-linking for protein interaction analysis of non-tagged integrin beta1. *J. Biomed. Biotechnol.*, 2010:927585, 2010.

[136] E. L. Rudashevskaya, R. Sacco, K. Kratochwill, M. L. Huber, M. Gstaiger, G. Superti-Furga, and K. L. Bennett. A method to resolve the composition of heterogeneous affinity-purified protein complexes assembled around a common protein by chemical cross-linking, gel electrophoresis and mass spectrometry. *Nat Protoc*, 8(1):75–97, Jan 2013.

[137] F. Stengel, R. Aebersold, and C. V. Robinson. Joining forces: integrating proteomics and cross-linking with the mass spectrometry of intact complexes. *Mol. Cell Proteomics*, 11(3):R111.014027, Mar 2012.

[138] D. J. Taatjes. The human Mediator complex: a versatile, genome-wide regulator of transcription. *Trends Biochem. Sci.*, 35(6):315–322, Jun 2010.

[139] N. J. Krogan, G. Cagney, H. Yu, G. Zhong, X. Guo, A. Ignatchenko, J. Li, S. Pu, N. Datta, A. P. Tikuisis, T. Punna, J. M. Peregrin-Alvarez, M. Shales, X. Zhang, M. Davey, M. D. Robinson, A. Paccanaro, J. E. Bray, A. Sheung, B. Beattie, D. P. Richards, V. Canadien, A. Lalev, F. Mena, P. Wong, A. Starostine, M. M. Canete, J. Vlasblom, S. Wu, C. Orsi, S. R. Collins, S. Chandran, R. Haw, J. J. Rilstone, K. Gandi, N. J. Thompson, G. Musso, P. St Onge, S. Ghanny, M. H. Lam, G. Butland, A. M. Altaf-Ul, S. Kanaya, A. Shilatifard, E. O'Shea, J. S. Weissman, C. J. Ingles, T. R. Hughes, J. Parkinson, M. Gerstein, S. J. Wodak, A. Emili, and J. F. Greenblatt. Global landscape of protein complexes in the yeast Saccharomyces cerevisiae. *Nature*, 440(7084):637–643, Mar 2006.

[140] A. J. Koster and J. Klumperman. Electron microscopy in cell biology: integrating structure and function. *Nat. Rev. Mol. Cell Biol.*, Suppl:6–10, Sep 2003.

[141] F. Plitzko Kourkoutis and W. J. M. Baumeister. Electron Microscopy of Biological Materials at the Nanometer Scale. *ANNUAL REVIEW OF MATERIALS RESEARCH.*, 42:33–58, August 2012.

[142] W. Kukulski, M. Schorb, S. Welsch, A. Picco, M. Kaksonen, and J. A. Briggs. Correlated fluorescence and 3D electron microscopy with high sensitivity and spatial precision. *J. Cell Biol.*, 192(1):111–119, Jan 2011.

[143] W. Baumeister. Electron tomography: towards visualizing the molecular organization of the cytoplasm. *Curr. Opin. Struct. Biol.*, 12(5):679–684, Oct 2002.

[144] D. Vanhecke, S. Asano, Z. Kochovski, R. Fernandez-Busnadiego, N. Schrod, W. Baumeister, and V. Lucic. Cryo-electron tomography: methodology, developments and biological applications. *J Microsc*, 242(3):221–227, Jun 2011.

[145] S. Nickell, C. Kofler, A. P. Leis, and W. Baumeister. A visual approach to proteomics. *Nat. Rev. Mol. Cell Biol.*, 7(3):225–230, Mar 2006.

[146] A. Leis, B. Rockel, L. Andrees, and W. Baumeister. Visualizing cells at the nanoscale. *Trends Biochem. Sci.*, 34(2):60–70, Feb 2009.

[147] J. Bohm, A. S. Frangakis, R. Hegerl, S. Nickell, D. Typke, and W. Baumeister. Toward detecting and identifying macromolecules in a cellular context: template matching applied to electron tomograms. *Proc. Natl. Acad. Sci. U.S.A.*, 97(26):14245–14250, Dec 2000.

[148] M. Beck, J. A. Malmstrom, V. Lange, A. Schmidt, E. W. Deutsch, and R. Aebersold. Visual proteomics of the human pathogen Leptospira interrogans. *Nat. Methods*, 6(11):817–823, Nov 2009.

[149] C. Bouchet-Marquis, M. Pagratis, R. Kirmse, and A. Hoenger. Metallothionein as a clonable high-density marker for cryo-electron microscopy. *J. Struct. Biol.*, 177(1):119–127, Jan 2012.

[150] S. Kemmerling, S. A. Arnold, B. A. Bircher, N. Sauter, C. Escobedo, G. Dernick, A. Hierlemann, H. Stahlberg, and T. Braun. Single-cell lysis for visual analysis by electron microscopy. *J. Struct. Biol.*, 183(3):467–473, Sep 2013.

[151] S. Kemmerling, J. Ziegler, G. Schweighauser, S. A. Arnold, D. Giss, S. A. Mueller, P. Ringler, K. N. Goldie, N. Goedecke, A. Hierlemann, H. Stahlberg, A. Engel, and T. Braun. Connecting mu-fluidics to electron microscopy. *J. Struct. Biol.*, 177(1):128–134, Jan 2012.

[152] M. Lim and K. J. Rothschild. Photocleavage-based affinity purification and printing of cell-free expressed proteins: application to proteome microarrays. *Anal. Biochem.*, 383(1):103–115, Dec 2008.

[153] G. C. Lander, S. M. Stagg, N. R. Voss, A. Cheng, D. Fellmann, J. Pulokas, C. Yoshioka, C. Irving, A. Mulder, P. W. Lau, D. Lyumkis, C. S. Potter, and B. Carragher. Appion: an integrated, database-driven pipeline to facilitate EM image processing. *J. Struct. Biol.*, 166(1):95–102, Apr 2009.

[154] A. Bremer, C. Henn, A. Engel, W. Baumeister, and U. Aebi. Has negative staining still a place in biomacromolecular electron microscopy? *Ultramicroscopy*, 46(1-4):85–111, Oct 1992.

[155] P. C. da Fonseca and E. P. Morris. Structure of the human 26S proteasome: subunit radial displacements open the gate into the proteolytic core. *J. Biol. Chem.*, 283(34):23305–23314, Aug 2008.

[156] J. Martin-Benito, J. Grantham, J. Boskovic, K. I. Brackley, J. L. Carrascosa, K. R. Willison, and J. M. Valpuesta. The inter-ring arrangement of the cytosolic chaperonin CCT. *EMBO Rep.*, 8(3):252–257, Mar 2007.

[157] Q. Ramadan and M. A. Gijs. Simultaneous sample washing and concentration using a "trapping-and-releasing" mechanism of magnetic beads on a microfluidic chip. *Analyst*, 136(6):1157–1166, Mar 2011.

[158] F. C. O'Mahony, J. Nanda, A. Laird, P. Mullen, H. Caldwell, I. M. Overton, L. Eory, M. O'Donnell, D. Faratian, T. Powles, D. J. Harrison, and G. D. Stewart. The use of reverse phase protein arrays (RPPA) to explore protein expression variation within individual renal cell cancers. *J Vis Exp*, (71), 2013.

[159] E. B. Tikhonova, A. S. Ethayathulla, Y. Su, P. Hariharan, S. Xie, and L. Guan. A transcription blocker isolated from a designed repeat protein combinatorial library by in vivo functional screen. *Sci Rep*, 5:8070, 2015.

[160] V. Cortez-Retamozo, N. Backmann, P. D. Senter, U. Wernery, P. De Baetselier, S. Muyldermans, and H. Revets. Efficient cancer therapy with a nanobody-based conjugate. *Cancer Res.*, 64(8):2853–2857, Apr 2004.

[161] K. E. Conrath, M. Lauwereys, M. Galleni, A. Matagne, J. M. Frere, J. Kinne, L. Wyns, and S. Muyldermans. Beta-lactamase inhibitors derived from single-domain antibody fragments elicited in the camelidae. *Antimicrob. Agents Chemother.*, 45(10):2807–2812, Oct 2001.

[162] A. D. Ellington and J. W. Szostak. In vitro selection of RNA molecules that bind specific ligands. *Nature*, 346(6287):818–822, Aug 1990.

[163] P. Broz and D. M. Monack. Molecular mechanisms of inflammasome activation during microbial infections. *Immunol. Rev.*, 243(1):174–190, Sep 2011.

[164] A. Engel. Assessing biological samples with scanning probes. In Astrid Gräslund, Rudolf Rigler, and Jerker Widengren, editors, *Single Molecule Spectroscopy in Chemistry, Physics and Biology*, volume 96 of *Springer Series in Chemical Physics*, pages 417–431. Springer Berlin Heidelberg, 2010. ISBN 978-3-642-02596-9. doi: 10.1007/978-3-642-02597-6_21. URL http://dx.doi.org/10.1007/978-3-642-02597-6_21.

[165] X. Yu, Y. Hiromasa, H. Tsen, J. K. Stoops, T. E. Roche, and Z. H. Zhou. Structures of the human pyruvate dehydrogenase complex cores: a highly conserved catalytic center with flexible N-terminal domains. *Structure*, 16 (1):104–114, Jan 2008.

[166] Ji Zhou, Zhihong He, Yu Ma, and Shikui Dong. Study of light-absorbing crystal birefringence and electrical modulation mechanisms for coupled thermal-optical effects. *Appl. Opt.*, 53(27):6243–6255, Sep 2014. doi: 10.1364/AO.53.006243. URL http://ao.osa.org/abstract.cfm?URI=ao-53-27-6243.

[167] R. Kitamura, L. Pilon, and M. Jonasz. Optical constants of silica glass from extreme ultraviolet to far infrared at near room temperature. *Appl Opt*, 46 (33):8118–8133, Nov 2007.

[168] G. M. Hale and M. R. Querry. Optical Constants of Water in the 200-nm to 200-microm Wavelength Region. *Appl Opt*, 12(3):555–563, Mar 1973.